# Using Video Annotation Software to Develop Student Self-Regulated Learning

# Using Video Annotation Software to Develop Student Self-Regulated Learning

Final report 2015

University of South Australia
The University of Sydney
The University of New South Wales
University of Edinburgh
University of British Columbia

Project team and report authors:

Professor Shane Dawson
Dr Abelardo Pardo
Dr Negin Mirriahi
Professor Dragan Gasevic
Professor Alan Kingstone

< http://ltudl-shib1.cw.unisa.edu.au/wordpress/>

Support for the production of this report has been provided by the Australian Government Office for Learning and Teaching. The views expressed in this report do not necessarily reflect the views of the Australian Government Office for Learning and Teaching.



2015

| | | |
|---|---|---|
| ISBN | PRINT | 978-1-76028-547-0 |
| ISBN | PDF | 978-1-76028-548-7 |
| ISBN | DOCX | 978-1-76028-549-4 |

# Acknowledgements

The Project Team would like to acknowledge the expertise and commitment of the many individuals that so generously provided their insights and reflections on this research. In particular we would like to thank the members of the national and international reference group for their time and assistance in guiding the development of the video annotation tool, the analyses and the final preparation of this report.

The project team would also like to acknowledge the prior work on the Collaborative Lecture Annotation System (CLAS) undertaken by Tom Foulsham, Evan Risko and Allan Kingstone. The original CLAS tool was developed at the University of British Columbia and this project builds off this initial software. The team very much acknowledges the excellent technical work undertaken by Thomas Dang (Arts ISIT, UBC) and An Zhao (TIU, UniSA). The team would also like to thank Professor Fred Cutler (UBC) for his collaborations and ongoing support of the CLAS tool.

# List of acronyms used

| | |
|---|---|
| AAF | Australian Access Federation |
| Arts ISIT | Faculty of Arts, Instructional Support and Information Technology, University of British Columbia |
| BYOD | Bring Your Own Device |
| CLAS | Collaborative Lecture Annotation System |
| ICT | Information and Communication Technologies |
| JOL | Judgement of learning |
| LA | Learning Analytics |
| LMS | Learning Management System |
| OLT | Office for Learning and Teaching (Australian Government) |
| OVAL | Online Video Annotation for Learning (revised development of the CLAS software) |
| SOLAR | Society for Learning Analytics Research |
| SIS | Student Information System |
| SRL | Self-regulated learning |
| TIU | Teaching Innovation Unit, University of South Australia |
| UBC | University of British Columbia, Vancouver, Canada |
| UniSA | University of South Australia |
| UNSW | University of New South Wales |
| USyd | The University of Sydney |

# Executive summary

The adoption of video for learning and teaching purposes has shown a dramatic increase over the past decade. As educational institutions increasingly embrace blended and online learning modalities there has been a corresponding reliance on the use of video to provide disciplinary content and exemplars as well as extending opportunities for learner self-reflections. Although video has long been an integral resource for educators, it is only recently that the availability and affordability of computing/ mobile devices alongside the marked increase in connectivity that has generated such ease of accessibility to these learning resources. Any student can now create, or capture and distribute video via a myriad of technologies from mobile phones to tablets or notebook devices – and let's not forget the more traditional video camera. As such, we are now witness to a wealth of multi-media development such as automated lecture capture, mobile video capture, and live streaming of events though mobile applications such as periscope[1] and the establishment of university video production units to generate high quality learning materials. Despite the growth of video in education we know little about how such resources are effectively implemented to improve learning and promote skills and proficiency in self-regulated learning (Szpunar, Jing, & Schacter, 2014).

## Project aims

This collaborative innovation and development project sought to address this deficit by investigating the application of video and video annotations in higher education courses in order to improve student self-regulated learning (SRL) skills. The project aimed to leverage the primary products of existing pedagogical practice (e.g. lecture capture) in a manner that fosters student SRL and reflection on teaching. The project objectives included:

- investigate learner and teacher interactions with recorded video content to promote cognitive engagement and develop recommendations for enhancing video annotation software (case study 1, 2 and 3);
- establish lead indicators for assessing student self-regulated learning – skills and attributes associated with 21st Century literacies (case study 1 and 2);
- develop and establish data visualisations and feedback processes for fostering student SRL proficiency (case study 1 and 2); and
- develop analytics and visualisations to empower faculty regarding decisions about their teaching and student learning progression (case study 1, 2 and 3).

[1] https://www.periscope.tv/about

## Approach

To address the project objectives a series of case studies were implemented across a range of pedagogical contexts and disciplines. This project report provides details on the outcomes and findings from three specific case studies. The first case study based at the University of British Columbia (UBC, Vancouver, Canada) involved a longitudinal analysis within the School of Music. This particular case study provided detailed insight into student uptake and self-reflective annotations. The second case study was based in the Faculty of Engineering at The University of Sydney (USyd) and aimed to address the use of video for conceptual understanding of disciplinary content for lecture preparation (flipped learning) within a large class setting. The case study also included a quantitative approach for assessing student learning strategies and self-regulated learning proficiency. The case study provides insight into the association between the adoption of video annotations as a learning strategy and the observed improvement in student self-efficacy. The third case study involved a teacher training program at the University of New South Wales (UNSW). The foundations course was targeted to identify how faculty can use such tools to aid the design and delivery of blended learning formats. The case study provided insights into discrete learner profiles based on student interactions with the implemented video annotation tool.

The successful completion of the project was reliant on extensive redevelopment of the video annotation platform called CLAS - Collaborative Lecture Annotation System. This software was first co-developed by members of the project team and other international researchers (see Dawson, Macfadyen, Risko, Foulsham, & Kingstone, 2012; Risko, Foulsham, Dawson, & Kingstone, 2013). Following the re-development of CLAS and the completion of this Office for Learning and Teaching (OLT) project the software has been re-named to OVAL (Online Video Annotation for Learning). This re-naming emphasises the key annotation feature and distinguishing characteristic of the software. The additional development for OVAL included a suite of learning analytics that provide detailed data regarding student viewing and play event interactions as well as the number, timing and types of annotations completed. The software includes the ability for teachers to provide automated feedback (points) relating to the embedded video that corresponds to individual student's judgement of learning (JOL) and accuracy of their written summary/ annotations. As demonstrated through the various case studies, the analytics provide insight into student proficiency in SRL as well as acting as a potential early-alert system for students requiring additional learning support (Dawson et al., 2012; Mirriahi & Dawson, 2013; Risko et al., 2013).

## Summary of outcomes

The project has achieved its stated objectives and outcomes. The study of how students interact with the video annotation tool across various learning scenarios has established a set of lead indicators for identifying proficiency of self-regulated learning. These indicators have then been distilled to a set of associations between the use of the tool and academic performance. Additionally, the data collected has allowed for the identification of a set of discrete learner profiles to aid development of automated yet personalised instructive feedback. The combination of these profiles with the aforementioned indicators offers a solid base to connect personalisation and academic improvement that is both scalable and instructive. In the area of tool development, users can now embed videos directly from YouTube thus removing the need for institutionally supported media repositories. This enhancement also adds an extra level of flexibility as instructors can quickly include any publicly available video hosted in YouTube to the learning environment. Furthermore, the underlying design structure provides the opportunity to integrate with other proprietary media tools such as Kaltura, or Helix Media repository. The software has its authentication layer integrated with the Australian Access Federation (AAF) allowing for seamless authentication credentials management with any higher education institution in Australia. As noted above the tool includes a suite of detailed learning analytics reports to facilitate personalised feedback and promote student SRL. The design of CLAS includes an adaptive layout to ensure compatible with multiple devices including mobile phones.

## Recommendations for practice

The project execution and more precisely the experimental studies have allowed the team of researchers to identify a set of key recommendations for practice. The deployment of video annotation tasks has been identified as a powerful enhancement of audio visual content with a high potential for students to reflect in content. However, this deployment requires the proper scaffolding within the learning experience. Students should be made aware of the purpose and potential of the video annotation tasks. Specific measures towards this end should be part of the learning experience. The use of video annotation should be part of a more comprehensive strategy to adopt a deep learning approach. The way in which students adopt the video annotation tool can be aligned to discrete user profiles. Early identification of these profiles can facilitate the development of personalised feedback and promote skills in SRL. Video annotation can provide a larger impact when its use is personalised for different student profiles.

# Table of contents

Acknowledgements.....3

List of acronyms used.....4

Executive summary.....5

Project aims.....5

Approach.....6

Summary of outcomes.....6

Recommendations for practice.....7

Table of contents.....8

Tables and figures.....10

Tables.....10

Figures.....10

1 Introduction.....12

2 Literature review.....13

2.1.1 Online Video Annotation for Learning (OVAL).....15

2.1.2 Self-regulated learning.....16

2.2 Project goals.....18

3 Case Study 1: Performing Arts Courses.....19

3.1.1 Study Context.....19

3.2 Volume of Use of CLAS.....21

3.3 Content Analysis of Video Annotations.....24

4 Case Study 2: First Year Engineering Course.....26

4.1 Integration of CLAS in the course design.....26

4.2 Experimental Results.....27

5 Case Study 3: Professional Development Program.....30

5.1 Description of how CLAS was integrated into the program design.....30

5.2 Engagement Patterns with CLAS ........ 31

5.2.1 Covariates as predictors of class membership ........ 32

6 Discussion and Findings ........ 34

7 References ........ 37

Appendix A ........ 40

Appendix B ........ 41

Appendix C ........ 43

Appendix D ........ 52

Appendix E ........ 56

Appendix F ........ 59

# Tables and figures

## Tables

Table 1: Descriptive statistics of the variables of the study ..................................................28
Table 2: Association between CLAS use and midterm marks...............................................28
Table 3: Association between CLAS and midterm marks after controlling for the five MSLQ scales ........................................................................................................................29
Table 4: Association between the use of CLAS and the midterm after controlling for the SPQ scales ........................................................................................................................29
Table 5. Comparison of the operations between Course 1 and Course 2..............................44
Table 6. Comparison of the learning products and learning strategy between Course 1 and Course 2 ....................................................................................................................45
Table 7. Comparison of the operations between Course 2 and Course 3..............................46
Table 8. Comparison of the learning products and strategy between Course 2 and Course 3 ...............................................................................................................................48
Table 9. Comparison of the operations between Course 2 and Course 4..............................49
Table 10. Comparison of the learning products and operations between Course 2 and Course 4 ..........................................................................................................................51
Table 11. Coding scheme used for content analysis of the specificity of reflections.............53
Table 12. Inferential statistics for the model fit assessment of the effects of instructional condition and experience on the specificity of reflective activities with video annotation ...54

## Figures

Figure 1: Annotations (point based and text) aligned with the video time stamp ................16
Figure 2: Sites of convergence/ divergence from individual student annotations ................16
Figure 3: video based annotations. These forms of annotations have been adopted for practitioner feedback and review in medical and performance based education (e.g. Theatre and film, music)..........................................................................................................16
Figure 4: The courses included in the study and student groups formed based on enrollment in individual courses ......................................................................................................20
Figure 5: Comparisons of the counts of the students' use of CLAS and time spent on interacting with videos in the four courses in case study 1. ..................................................22
Figure 6: Effects of instructional conditions and experience with video annotation on the counts of general annotations and specificity of self-evaluations (i.e., reflections)..............25
Figure 7: Latent classes of participants across time. .............................................................32
Figure 8: Plot of class membership as influenced by the participants' current position at the university..................................................................................................................33
Figure 9: AAF Initial login screen..........................................................................................41
Figure 10: Institutional login screen ......................................................................................41
Figure 11: User content in the professional development course ........................................42

Figure 12. Solutions for fixed effects (instructional condition and experience) on the specificity of reflective activities with video annotation........................................................55
Figure 13 Example of Activity using CLAS..................................................................56
Figure 14 Video activity in CLAS with General Comments and Annotations ..........................57
Figure 15 Dialogue presented upon general comment submission.......................................58

# 1 Introduction

Technologies have become an indispensable part of teaching and learning practice. Over the past decade we have witnessed massive uptake in student ownership of laptop and mobile devices. This explosion of ownership of computing technologies has been well encapsulated in the BYOD (Bring Your Own Device) movement in education. Internet connectivity is also an essential tool for contemporary education. Along with the proliferation of coffee shops, there has been a comparable growth in access to Wi-Fi. The increased ownership of computers, the sophistication of technologies and ease of access to the internet have combined to significantly influence how we undertake teaching and learning practice.

Central to this change has been the implementation of enterprise software (e.g. learning management systems – LMS) to facilitate access to course content and engage with peers and instructors. Essentially, the hardware and software have enabled the easy development and dissemination of rich multi-media such as simulations and video. While video has long been an integral resource for educators, it is only recently that these visual resources can be so easily developed and accessed. For instance, any student can now create, or capture and distribute video via a myriad of technologies from mobile phones to tablets or notebook devices – and let's not forget the more traditional video camera. Despite the extensive reliance on video in education we know little about how such resources are effectively used to improve learning and promote skills and proficiency in self-regulated learning (Szpunar, Jing & Schacter, 2014).

This collaborative innovation and development project aimed to address this challenge by leveraging the primary products of existing pedagogical practice (lecture content) in a manner that fosters student self-regulated learning (SRL) and reflection. The project first extended the development of the video annotation software CLAS (Collaborative Lecture Annotation System) (Dawson et al., 2012; Risko et al., 2013). These technical developments resulted in open and accessible software integrated into YouTube. The software was renamed as OVAL – Online Video Annotation for Learning and three case studies were undertaken to provide insights and evidence into how learner video annotations can be applied to positively impact on student self-regulated learning and academic performance.

The first case study aimed to investigate the effects of instructional conditions on the adoption and sustained use of the video annotation tool. Four undergraduate courses in the performing arts discipline were considered. Students recorded their performances in video and then annotated the videos with OVAL. The objective of the study was to measure the use of OVAL throughout two consecutive courses when the annotations were graded vs non-graded.

The second study case examined how video annotations are adopted within a first year engineering course drawing on a flipped learning curriculum design strategy. The objective was to promote social engagement between students through the integration of video content and associated annotations. The case study provided user interaction visualisations to students so that they were more aware of their progress. The trace data from student interactions with the CLAS tool, and reported learning approaches and motivations were used to identify new methods for assessing self-regulating learning.

The third case study was carried out in the context of a professional development program at The University of New South Wales (UNSW). Video annotation was used to identify exemplary blended and online approaches to course design and delivery. Annotations were shared among the students in the cohort to foster self-reflection and self-assessment. Similar to case study 2, the course followed a flipped learning strategy and the video annotations were part of the preparation activities before attending a face-to-face session. The objective of the study was to use video annotation to promote self-regulated learning and detect the most common usage patterns in this context.

The results from these three case studies have led to the identification of multiple findings that are important for contemporary learning and teaching practice.

# 2 Literature review

The Australian higher education landscape is undergoing significant change in response to government policy, increased market pressure as result of the globalisation of education and momentum associated with disruptive technologies such as Massive Open Online Courses - MOOCs and learning analytics. The sector has set clear targets around increasing university participation and graduation rates alongside calls for improving the student experience and developing graduates with the necessary 21$^{st}$ Century literacies for “full participation in society and the economy” (Commonwealth of Australia, 2009, p. 7). Even at a simplistic level, these initiatives have placed greater pressure on universities to establish scalable, cost effective and high quality teaching practices that foster 21$^{st}$ Century skills and attributes regardless of discipline, class size and modality.

The rhetoric for embracing technology enhanced learning is far easier than establishing broad scale adoption into daily practice. To date the adoption of educational technology has been predominantly via an institutional Learning Management System. Although the LMS provides a solid platform to deliver online and mixed mode education, it also promotes a more transmissive pedagogy – the dissemination of disciplinary knowledge through sharing of content files. That is not to say that the LMS is not used in collaborative and pedagogically rich designs. More so, that the adoption process has tended to mirror a lecture centric model into an online world and that innovative applications of educational technologies are to date sporadic at best (Conole & Alevizou, 2010; Cummins, Brown & Sayers, 2007). Given there is scant indication that this pedagogical tipping point from lecture centric to social learning will be reached in the near future (Conole & Alevizou, 2010), alternate approaches for leveraging the products of lecture based instruction to better engage our students in more socially oriented learning activities are necessary. However, a simple change in practice or leveraging the social affordances of new technologies alone will not result in broad scale sustained pedagogical transformations.

Any significant and sustained shift in pedagogical practice must be empirically driven. In this context, the field of learning analytics provides much rigour for demonstrating the impact of implemented teaching and learning approaches. Learning analytics draws on user digital trace data derived from student interactions with various information and communication technologies (ICTs). Essentially, these trace or log data are ‘mined’ and analysed to identify patterns of learning behaviour that can be used to inform education practice (Greller & Drachsler, 2012; Pardo & Kloos, 2011). While these approaches for analysing student data have been in practice for many decades, learning analytics differentiates itself by apply a strong theoretical (learning theory) foundation to the findings. The Society for Learning Analytics Research (SoLAR) defines learning analytics as the “measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (Siemens & Long, 2011).

Despite the relative immaturity of the field, learning analytics is rapidly building momentum towards broad-scale institutional adoption across the education sector (Colvin et al., 2015). This finding was well noted in the recent OLT strategic commissioned report on learning analytics in Australia. In short, all Australian universities were in the process of implementing a form of learning analytics to address areas such as student retention and academic performance. In so doing institutions were also seeking to extend their technical capacity and infrastructure to capture student data and provide easier reporting mechanisms. Although, LMS have for some time provided a suite of user tracking data, the levels of analyses presented are poorly understood and disconnected from the learning design. These forms of analysis and data visualisation do little to inform educators of actual student understanding and therefore provide limited insight into the learning process (Lockyer, Heathcote & Dawson, 2013). A more advanced analysis is required to shed light into the learning processes so that both instructors and students are aware of and identify the impact of the learning experience. It is at this junction of establishing fine grained, automated and scalable learning analytics while leveraging the artefacts of current dominant practice that this collaborative project aimed to address. As detailed in the following section, the project builds upon a novel educational technology that shifts the focus of teaching practice from transmission of content towards social and personalised learning. A key objective of the project was to provide a scalable means for enhancing the student experience through establishment of fine-grained learning analytics. The project was grounded in self-regulated learning research (SRL) (Pintrich & Zusho, 2007; Winne, 2011; Zimmerman, 2008) and as such, the associated analytics and visualisations serve to promote and assess student SRL during their engagement with video and video annotations.

### 2.1.1 Online Video Annotation for Learning (OVAL)

A core objective for the project was to investigate the application of video annotation software as a means for promoting teaching and learning practices that are associated with the development of student SRL. The video annotation software, CLAS, was first developed by the project team at the University of British Columbia (UBC) in 2011 (Risko et al., 2013). At this time the CLAS software was a web-based annotation tool for students and teachers to annotate videos/audio (lecture capture, student performance videos in music or theatre) in the same way they highlight important sections within a text medium. Students and instructors used CLAS to interact directly with any uploaded multi-media content (e.g. podcasts, lecture capture, PowerPoint). For example as a student watches an uploaded lecture within the CLAS software where they can annotate, on a timeline, the important points and concepts (Figure 1). These annotations can be shared, included as part of a peer review process or available for instructor feedback in one collaborative space regardless of class size or discipline.

While the software shares functionality with other annotation tools such as videoANT[2], it does provide a unique set of features. These include the capacity for point-based annotations (Figure 1); visualisation of areas of student convergence and divergence (Figure 2), and the inclusion of video-to-video annotations (Figure 3). The capacity for learners and instructors to review areas of convergence and divergence in CLAS (e.g. uploaded lecture) allows for rapid identification of student learning and the effectiveness of the lecture to convey key points (Risko et al., 2013). For example, students can assess their own interpretation of important points against those of their peers along the video timeline. Teachers can review the sites of convergence to ensure key concepts were covered. Annotations that lie outside the areas of convergence can be quickly identified and accessed to check for any conceptual misinterpretations. In this way, visualisation of the captured student interactions can inform decisions regarding future teaching strategies and practices in order to address any identified gaps in the learning process or student understanding (Dawson et al., 2012). CLAS is a scalable tool that has clear potential for leveraging video content (lecture capture) to engage students in more socially oriented learning practices. The prototype CLAS tool demonstrated much potential; however, further enhancements were required to include features such as judgement of learning (JOL), instructor/peer feedback and learning analytics. This project aimed to advance the software development (as open source) to include features that would promote student reflection and development of self-regulated learning.

[2] http://ant.umn.edu/

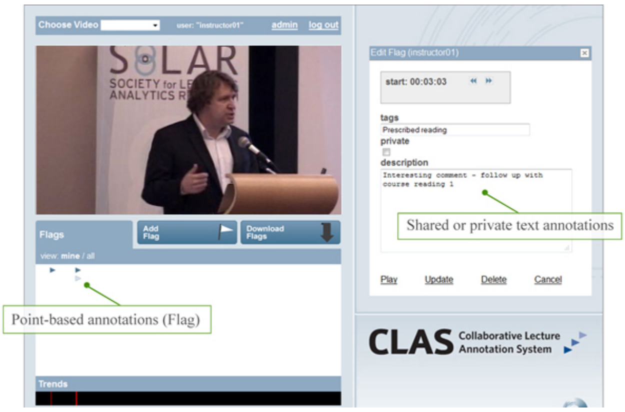



**Figure 1:** Annotations (point based and text) aligned with the video time stamp

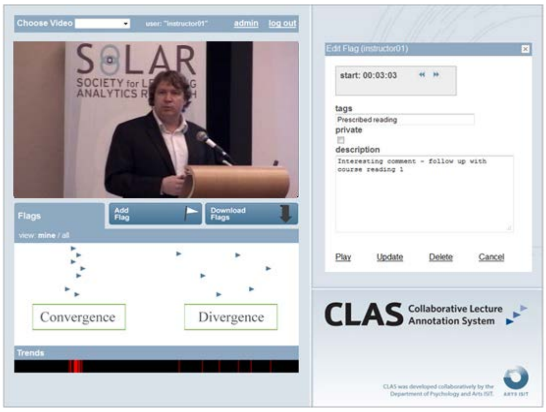



**Figure 2:** Sites of convergence/ divergence from individual student annotations

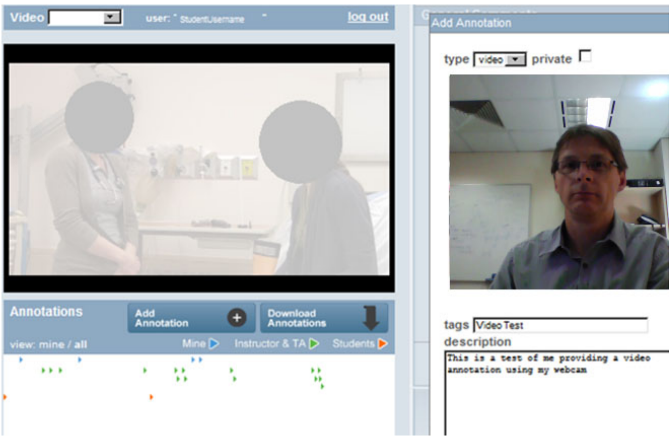



**Figure 3:** video based annotations. These forms of annotations have been adopted for practitioner feedback and review in medical and performance based education (e.g. Theatre and film, music).

### 2.1.2 Self-regulated learning

The widespread adoption and pace of technological developments, global markets, and the increasing focus towards a knowledge economy have coalesced to play a greater role in determining an individual's future access to and productive participation in local, global and virtual societies. The skills and attributes graduates now require for productive participation have been encapsulated in the discourse surrounding 21$^{st}$ Century capacities. These related skills and attributes centre on developing self-regulated learners. Student proficiency in self-regulated learning (SRL) has been demonstrated to be strongly correlated with academic achievement (Eccles & Wigfield 2002). It is widely accepted that this stems from SRL relating to an individual's ability to accurately monitor and regulate their learning behaviours. However, SRL requires a well-developed level of self-motivation and meta-cognitive capacity – areas that have been demonstrated to be weak among undergraduate cohorts (Bjork, Dunlosky & Kornell, 2013).

Judgement of learning (JOL) is a commonly adopted process for both assessing and developing SRL. The JOL process involves a student making an assessment of how effectively they have understood a concept. For example, a student may be assigned a course reading and is then asked to rate their confidence in demonstrating understanding of the concepts and key points. This self-reported rating is correlated against overall future test performance. Numerous studies have demonstrated that students are often over-confident in their evaluation of their learning, which can lead to detrimental study patterns and behaviour, even withdrawal (Dunlosky & Rawson, 2012). However, developing an individual's JOL leads to more targeted study behaviours and improved academic performance (Kruger & Dinning, 1999). While JOL has commonly relied on self-report methodologies such as think aloud protocols and surveys, alternate options are emerging as a result of advances in learning analytics, machine learning and text analysis for example.

Due to the self-report nature, past JOL methods have inherent limitations such as poor recall, and biased responses (Richardson, 2004). More recently, alternate models attempting to address these issues seek to capture actual behaviour (in lieu of self-reports) from learner interactions with ICT based activities. For instance, recent studies demonstrate the capacity for student online interaction data to provide significant lead indicators of SRL; e.g., (Winne, 2010; Zimmerman, 2008). While activity logs can help address the limitations of self-reports, students' JOL can be improved by using reflection techniques. For instance, summarising text and generating keywords describing the concept (learning task and goal) have been effective for improving JOL and therefore overall academic performance (de Bruin, Thiede, Camp & Redford, 2011). These methods are further enhanced when the individual learner receives feedback on their inputs prior to undertaking a JOL (Rawson, Dunlosky & Thiede, 2000). Given the importance of user input and reflection through keywords and summary text, the CLAS software includes features that effectively leverage these processes for enhancing SRL. Hence, this project developed the capacity for teachers to embed keywords and summary comments to an uploaded video for enhancing student JOL and therefore SRL capacity.

In summary, this project re-purposed multimedia artefacts, such as lecture capture and course videos to engage students in cognitively challenging and socially oriented learning experiences through the use of a video annotation software. The CLAS software provided an effective resource to capture user viewing interactions and text annotations. These trace data analysed to determine the potential for video activity to both assess and facilitate development of student SRL. As the following case studies illustrate, the log data can be analysed to effectively identify user profiles or behaviour clusters related to SRL strategies. The combined suite of available learning analytics assists students and teachers in assessing learning progression, levels of engagement, and sites for improvement in a scalable and automated process.

## 2.2 Project goals

This innovation and development project addressed the need for educators to not only design activities consistent with the development of 21$^{st}$ Century learning skills but to also identify students requiring learning support early in their academic studies. As such the project aimed to:

1. research learner and teacher interactions with recorded video content to promote social and cognitive engagement and develop recommendations for enhancing the CLAS software;
    a. leverage the artefacts of a transmissive pedagogy (e.g. lectures) to promote more social learning opportunities;
    b. further develop and refine CLAS for mobile learning and cross platform compatibility;
2. provide faculty with direct feedback regarding the efficacy with which they are communicating the central points in their lectures to their students (Mirriahi & Dawson, 2013; Risko et al., 2013); and provide students with intuitive visualisations of their learning progress within the learning community as well as information for self-reflection (Janssen, Erkens & Kirschner, 2011; Jovanovic, Bagheri & Gasevic, 2015);
    a. development of data visualisations aligned with the associated learning activity (e.g. peer review, self-refection, practitioner review, or assessment);
    b. deployment of analytics and visualisations to empower faculty and students regarding decisions about their teaching and learning progression;
3. establish nuanced lead indicators for assessing student self-regulated learning –skills and attributes associated with 21$^{st}$ Century literacies;
    a. identify and validate new methods for assessing SRL from fine-grained, temporally ordered data derived from student learning activities;
    b. develop and establish data visualisations and automated feedback processes for fostering student SRL proficiency; and
4. develop international and national multi-campus longitudinal case studies related to the adoption of CLAS across a variety of disciplines and contexts with a view to establish analytics consistent with the pedagogical context.

# 3 Case Study 1: Performing Arts Courses

The first case study was conducted at the University of British Columbia, Vancouver Canada. The case study was longitudinal spanning several semesters of study and multiple courses. This component of the project aimed to investigate the effects of the course instructional conditions on student use and adoption of the CLAS software (objectives 1 and 2). Although existing research offers insights into the factors predicting students' intentions to accept a learning tool, much less is known about factors that affect actual adoption and sustained use and how such tools can be used to identify and promote self-regulated learning. Specifically, the study looked at the influence of course design, instructor and student feedback and the implemented assessment models on student adoption and use of CLAS. The CLAS tool was used for student's to reflect (self-reflection) on their individual and group musical performances. As the study was longitudinal there was an associated opportunity to examine how the course instructional conditions influenced an individual's long-term use.

### 3.1.1 Study Context

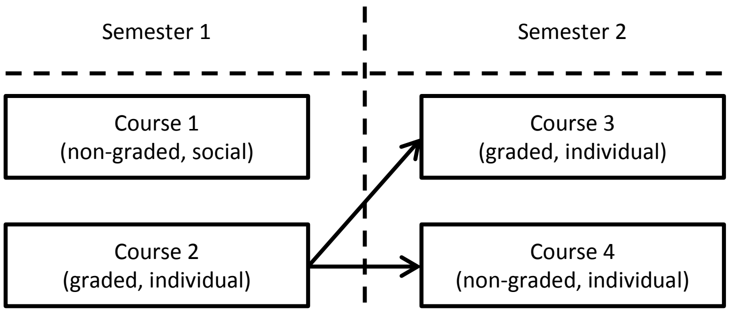
Four undergraduate courses based in the performing arts and offered over the 2012/2013 academic year were included in the study. The first two courses (Course 1 and Course 2) took place in the first semester of 2012. The other two courses (Course 3 and Course 4) took place in the second semester 2013[3]. In all four courses, the students were requested to perform, their performances were video recorded, and the recordings were available for viewing and annotating in the CLAS video annotation software. The annotations were used to reflect on the own performances. In Course 1, the students participated in group performances and thus, all the students enrolled into Course 1 viewed and annotated the same (three) video recordings. The students' annotations were not graded and the students received no formative feedback from the course instructor on their reflective annotations. The remaining three courses concentrated on the individual performance of the students. In Course 2, the students were requested to view the videos of their own four performances, reflect on their performances and enter their reflections as video annotations. For each video, students were requested to create time-stamped annotations and at least one general annotation per video summarizing their overall reflections on their performances. Of the four videos, the students received grades for their annotations on the three videos, while formative feedback was provided on their reflections of all the four videos.

Course 2 was a prerequisite for both Course 3 and Course 4 that were offered in the following semester. The requirements and instructional conditions for the use of CLAS in Course 3 were identical to those of Course 2 including graded general annotations. Course 3 had only one difference compared to Course 2 – it included an additional video recorded performance for students to view, annotate, and submit general reflective summaries in

[3] Note: the semester dates are based on the North American model whereby the academic year generally commences in September. Hence, the second semester for a program of study progresses into the following year.

CLAS. The instructional design of Course 4 was similar to Courses 2 and 3 in terms of the number of video recordings of the individual performances (i.e., four video recordings were accessible through CLAS). However, in contrast to Courses 2 and 3, the students did not receive a grade nor feedback on their general reflective annotations posted in CLAS. The relationship between the courses and the non-graded/graded general annotations are illustrated in Figure 4.

The students enrolled in Courses 1-4 were included in the study. The four courses had different numbers of enrollments: Course 1 (N=31), Course 2 (N=40), Course 3 (N=28), and Course 4 (N=20). Given that the study was conducted in the ecologically valid setting of a higher education institution, the students could enroll into any course, provided that they met the course prerequisites. This resulted in a total of 98 unique students who were enrolled into the four courses. The numbers of students enrolled in Courses 1-4 and the labels of derived subgroups are shown in Figure 4. The students who were only enrolled in Course 1 belong to group Course 1a, while the students who were enrolled in both Course 1 and Course 2 belong to two groups: Course 1b for the variables about their activity in Course 1 and Course 2b for the variables about their activity in Course 2. Similar grouping principles were used to group students enrolled in Course 2 and Course 3, and in Course 2 and Course 4. Only a minor overlap in the enrolment between Course 3 and Course 4 was noted (four students), a sample insufficient for statistical analysis.




Labels of the groups of the students created based on the courses they were enrolled in:

*Course 1a* – students who took course 1 but not Course 2 (n=23)

*Course 1b* – students who took Courses 1 and 2 (n=8)

*Course 2a* – students who took Course 2, but not course 1 (n=32)

*Course 2b* – students who took Courses 1 and 2 (n=8)

*Course 2c* – students who took Course 2, but not Course 3 (n=22)

*Course 2d* – students who took Course 2 and 3 (n=18)

*Course 2e* – students who took Course 2, but not Course 4 (n=29)

*Course 2f* – students who took Courses 2 and 4 (n=11)

*Course 3a* – students who took Course 3, but not Course 2 (n=10)

*Course 3b* – students who took Courses 3 and 2 (n=18)

*Course 4a* – students who took Course 4, but not Course 2 (n=9)

*Course 4b* – students who took Courses 4 and 2 (n=11)

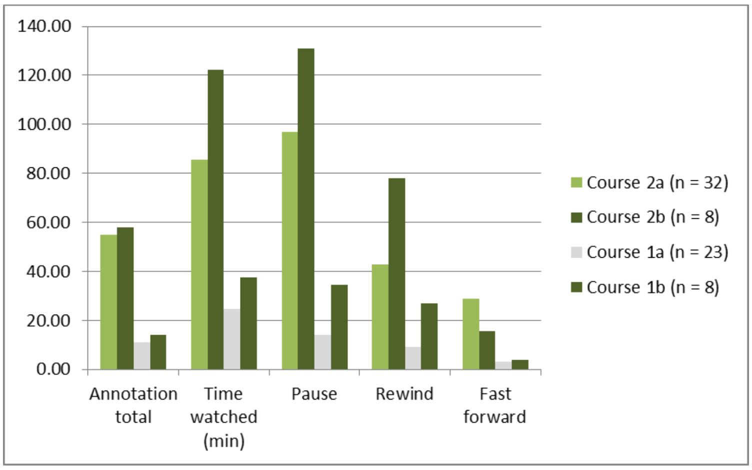
**Figure 4:** The courses included in the study and student groups formed based on enrollment in individual courses

## 3.2 Volume of Use of CLAS

Student enrolment in the four courses was used as the independent variable to study the effect of both independent and paired (repeated) samples. Counts of annotations created by the students were used as the primary dependent variable to measure the effect of the instructional scaffolding (graded or non-graded self-reflection annotations) on the student's sustained use of CLAS for any future course. The CLAS tool adopted for the study was designed to promote student engagement in the creation of video annotations and thus the quantity of annotations made reflects this primary focus. Several other dependent variables were also incorporated to understand the patterns of interactions with videos under different instructional conditions. These variables represented the occurrences of events recorded by CLAS in each course including the counts of pause, rewind and fast forward events. Similarly, the total time (minutes) when the play button was activated indicates the likely amount of time a user spent watching each video was included in the analyses.

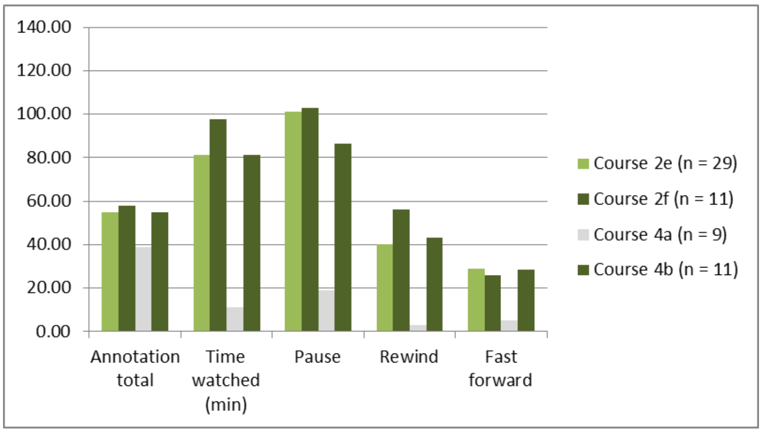
To evaluate differences in student learning strategies when using a video annotation tool for self-reflection, we created transition graphs based on the trace data of the recorded learning activities within CLAS (Hadwin, Nesbit, Jamieson-Noel, Code & Winne, 2007). Transitions graphs were created based on a contingency matrix where rows and columns accounted for all possible events. The rows represent the start of the event and the columns represent the end points of the transition edges. Originally, all the cells in the matrix had values of zero. If an event, A, was followed by an event B, number 1 was recorded in the cell representing the intersection of row A and column B. For each future appearance of this transition, the number in the cell was incremented by 1. In this way, we created weighted and directed transition graphs (networks). For this project, we distinguished between events based on the temporal parts of videos they were associated with. Specifically, we distinguished between time-stamped events (annotations, pause, rewind, and fast-forward) based on the quartiles of the videos they were associated with.

Non-parametric tests were applied to investigate the effects of the instructional conditions for the longitudinal study. To test for the effect of instructional conditions (i.e., independent variables), non-parametric tests for both independent (Mann-U Whitney test) and paired (Wilcoxon signed-rank test) samples were carried out in order to address the needs of the natural experiment the study was based upon. The detailed results of the statistical analyses are provided in Appendix C.




A)


B)




C)

**Figure 5:** Comparisons of the counts of the students' use of CLAS and time spent on interacting with videos in the four courses in case study 1.

Figure 5 shows the comparisons of the counts of the students' use of CLAS and time spent interacting with videos in the four courses in case study 1. Comparison between Courses 1 and 2 showed that although the median value of the number of videos annotated was consistent with the course requirements (three and four videos for Course 1 and Course 2, respectively, i.e., 25% higher requirements in Course 2), the time spent watching videos and the number of operations performed by the students when their general reflective annotations were assessed was by far greater than the increase of 25% in the requirements in Course 2. The learning strategy followed by the students in the assessed condition (Course 1 vs Course 2) exhibited a significantly higher level of metacognitive monitoring measured by the transition graph density than that of the students who were in the non-assessed condition only.

The comparison between the use of CLAS between Course 2 and Course 3 showed the effect of a student's prior experience with the tool and their overall quantity of activity. This was evident when the students used the tool in a subsequent course with the same assessed instructional conditions. A significant effect was observed for both the number of videos annotated and the total number of annotations created. While the increase in the number of videos annotated was anticipated, due to the increase in total videos in Course 3 (five videos) compared to Course 2 (four videos), the number of annotations created was much higher (almost doubled) than the increase in each student's performance video. Interestingly, the effect size of the instructional condition in Course 3 was larger for the students who had previously been enrolled in Course 2. This course (course 2) had the same instructional conditions in terms of the assessed general annotations. That is, students who used the tool for the second time (Course 3b) with the same assessed instructional condition had a tendency to create more annotations than the students who used the tool in the assessed instructional condition for the first time (Course 2d). The analysis did not reveal any significant differences in the level of metacognitive monitoring between the students in Course 2 (Mdn = .25 for both groups – Course 2c and Course 2d) and Course 3 (Mdn = .26 for group Course 3a and Mdn = .29 for group Course 3b) measured by the density of their transition graphs.

The comparison of the use of CLAS in Courses 2 and Course 4 investigated the effects of a student's previous assessed experience with CLAS on their future use of the tool in subsequent courses when there is a *non-assessed* instructional condition (i.e. general reflective annotations were not assessed). In this instance, there were no observed differences in the number of videos annotated. This was expected as the number of videos annotated by each student was consistent across the two courses (i.e. four in both of them). Likewise, there was no significant difference in the number of annotations created in the two courses. The comparison between students within Course 4 found those who previously had used CLAS in Course 2 (assessed instructional condition) made a greater number of annotations than those who used the tool in Course 4 for the first time. The numbers of pause, rewind, and fast forward events significantly declined with large effect sizes in Course 4 compared to those of Course 2. The metacognitive monitoring measured by the density of the transition graphs of students' learning strategy significantly dropped in Course 4 (Mdn = .06 for both groups – Course 4a and 4b) compared to Course 2 (Mdn = .25 for both groups – Course 2e and Course 2f) with a large effect size, regardless of previous experience with the tool.

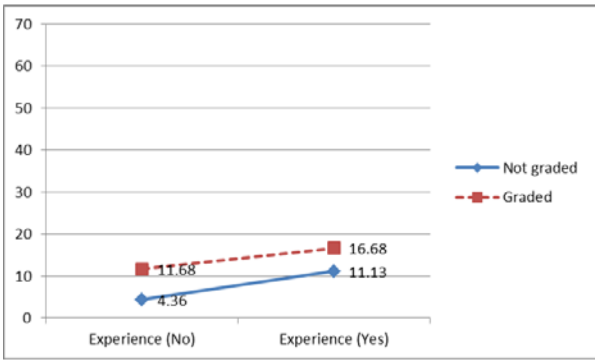
## 3.3 Content Analysis of Video Annotations

The first phase of this initial case study provided detailed data on the impact of instructional conditions on the quantity of interactions with the video annotation tool. The second phase of the case study queries whether the observed quantity of annotations also reflected any comparable increase in the quality of content associated with the annotations. A quantitative content analysis (Rourke & Anderson, 2004) was performed to assess the level of specificity of self-evaluations written in the students' reflections. The quantitative content analysis made use of a coding scheme that was adapted from Hulsman et al. (2009). The coding scheme consisted of three categories – observation, motive/effect, and goal. Further details concerning the coding scheme and the statistical analysis are provided in Appendix D.

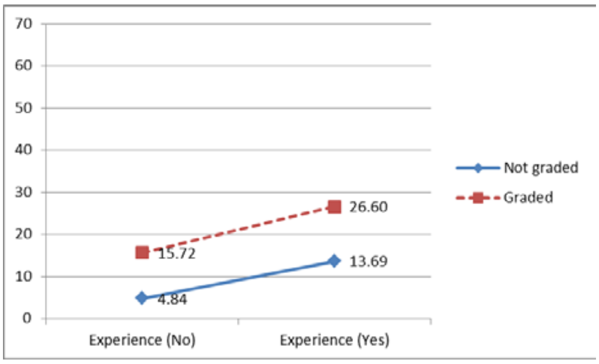
Multi-level mixed models were used for statistical analysis. Two independent variables were identified – the instructional condition and prior experience with the tool. These variables were categorical binary variables. The instructional condition variable represented whether the reflective learning experience with CLAS was both graded and formative feedback provided. The prior experience variable represented whether the students took a prior course in which reflective learning with CLAS was both graded and formative feedback provided. Prior experience was defined as students completing or not completing Course 2 prior to the time of data collection in Courses 3 and 4. The students in Course 1 and 2 did not have any prior experience. Counts of the appearances of the three self-evaluation specificity indicators – used for the assessment of students reflections (see Appendix D for details) – per student were used as the outcome variables. Counts of the total number of general annotations were also used to investigate the effects of instructional conditions and experience on the total volume of the activity performed with CLAS.

Figure 6 details the results of the estimated mean values based on the prior experience with CLAS in a graded condition with formative feedback and on instructional conditions (graded and formative feedback vs. non-graded and no formative feedback). Compared to the results of the total count values of the annotation creation across different conditions presented in Figure 5, the effect of prior experience with CLAS was only significant on the count values of general annotations created by Students (Figure 6a). Instructional conditions (i.e., graded along with formative feedback) had a statistical significance just under the 0.01 threshold. A similar pattern of statistical significance of prior experience only was found for the observation statements (Figure 6b), with a clear increase (though not statistically significant) in the number of observations when prior experience was combined with the graded condition accompanied with the feedback provision. The estimated mean value of the counts of effect/motive statements was much lower compared to the observation and goal categories (Figure 6c). Finally, prior experience, instructional condition with graded annotations and formative feedback, and the interaction of experience and instructional condition were significant for the counts of goal statements in the students' self-reflections

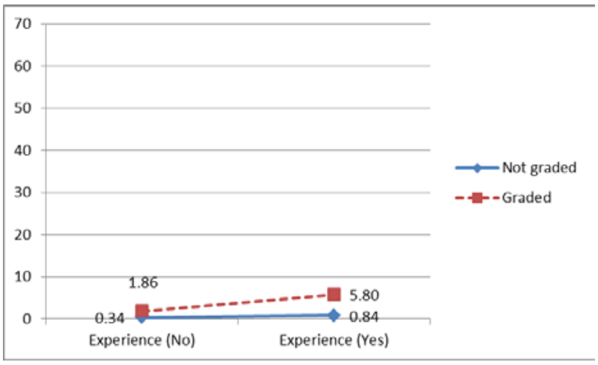
(Figure 6d). Further discussion of the practical recommendations related to this case study are presented in Section 6 - “discussion and findings” of this report.



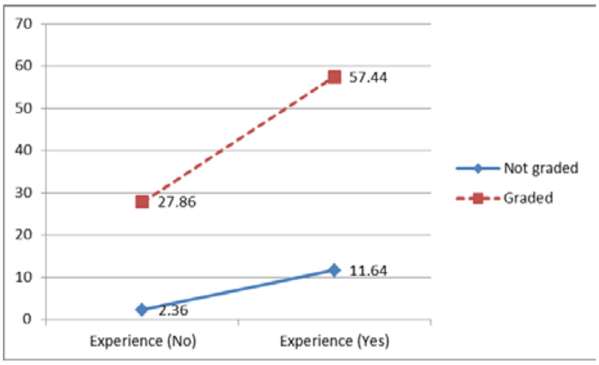

a) General annotations[1]




b) Observation[1]




c) Motive/effect[2,4]




d) Goal[2,3,4]

**Legend:** 1 – experience significant at $p<.01$; 2 – experience significant at $p<.001$; 3 – instructional condition significant at $p<0.01$; 4 – interaction of experience and instructional condition significant at $p<0.01$.

**Figure 6:** Effects of instructional conditions and experience with video annotation on the counts of general annotations and specificity of self-evaluations (i.e., reflections)

# 4 Case Study 2: First Year Engineering Course

This chapter describes the use of CLAS in the context of a first year engineering course as part of the Bachelor of Engineering program at The University of Sydney. The objective of this case study was related to the following project outcomes:

1. Promote social and cognitive engagement of students with video content. In this case the content was specifically recorded for the course and was designed to foster student engagement exchanging the summaries and annotations of several videos.
2. Deployment of analytics and visualisations to empower students regarding decisions about their teaching and learning progression.
3. Identify and validate new methods for assessing SRL from the collected data. The experiment established a relationship between student traits and the way they interact with video.

CLAS was used as the backbone of the course preparation strategy. This effectively required students to use the CLAS platform to share a summary of a set of pre-recorded videos explaining course content as if the lecturer was delivering it in a lecture hall. The introduction of the video summary was enhanced to pose questions and allow students to refine their submission. The objective of this functionality was to empower students regarding their decisions about their learning progression within the course and to promote student meta-cognitive skills. That is, the course design for integrating CLAS aimed to develop individual student's judgement of learning (JOL). The creation of annotations during the course activities offered students the possibility for sharing their insights about certain topics with their peers and review their relevance and accuracy of their contributions based on instructor feedback and questions. With this structure, students are able to reflect on concepts presented in a video, write a general summary and create annotations articulating their understanding of a particular course concept, and then check the accuracy of this summary through a reflective feedback process. These mechanisms contribute to the acquisition of self-regulated and judgement of learning skills.

## 4.1 Integration of CLAS in the course design

The learning outcomes of the engineering course included the configuration, building and testing of a computer application. The course draws on the embedded system Arduino[4] and uses a *hands on* approach to knowing the system and exposing the students to design tasks.

The course follows a conventional semester in the Australian higher education landscape with an overall duration of 13 weeks. The course was carried out in the second semester of a first year program. Enrolment was approximately 300 students. The large enrolment is in part due to the course acting as pre-requisite for multiple programs offered across the institution. Each week students would meet face to face over 2 hours for a *lecture* (plenary

[4] See the information in the home page of Arduino (http://www.arduino.cc)

sessions with all the students in a conventional lecture theatre), 2 hours in a tutorial (groups of 50 students with two tutors in the room) and 3 hours in a laboratory (groups of 70 students with three tutors in the room).

The course is delivered following a flipped learning model. This teaching approach required students to review and perform a set of activities prior to the lecture. The outcomes of the activities were further extended and discussed during the lecture session. The required activities mostly involved watching conceptual **videos** and formative assessment via a set of multiple choice questions. The deployment of CLAS in this environment was designed so that the application was tightly integrated with the current pedagogical framework. Towards this aim, each week in the course contained activities that required students to review a video using CLAS and then complete a series of formative assessment quizzes. To provide additional motivation students were awarded a summative score based on their completion of the multiple choice questions prior to each weekly lecture. This score simply formed part of the students overall summative assessment.

## 4.2 Experimental Results

The findings from the case study were published in the paper titled *Identifying Learning Strategies Associated with Active use of Video Annotation Software* presented at the International Conference on Learning Analytics and Knowledge[5]. The use of CLAS was combined with the scores obtained by the students in their midterm exam, and two conventional instruments to gauge student strategies: the Motivational and Self-Regulated Learning Questionnaire (MSLQ) (Pintrich & de Groot, 1990) and the Study Process Questionnaire (SPQ) (Biggs, Kember & Leung, 2001). These last instruments were deployed in the first session of the course. The independent variables used for the statistical analysis included the events captured by CLAS and the categories derived from MSLQ and SPQ. The dependent variable was the scores in the midterm exam. Table 1 details the descriptive statistics of the variables used in the study.

A first linear regression model computed considering solely the performance in the midterm exam and the use of CLAS is show in Table 2. The results shown in the table indicate significantly higher grades for the students who used CLAS compared to those who did not use CLAS. The variability of the students' mid-term grades explained by the model was less than 6%.

[5] Pardo, A., Mirriahi, N., Dawson, S., Zhao, Y., Zhao, A., & Gašević, D. (2015). Identifying learning strategies associated with active use of video annotation software. In Proceedings of the Fifth International Conference on Learning Analytics and Knowledge, Poughkeepsie, New York, USA, ACM, pp. 255-259.

**Table 1:** Descriptive statistics of the variables of the study

| Variables | Mean (SD) Entire sample (N=149) | Mean (SD) CLAS Users (N=41) | Mean (SD) CLAS Non-Users (N=108) |
|---|---|---|---|
| MID | 14.09 (3.61) | 15.43 (3.19) | 13.58 (3.64) |
| CLAS USE | 2.09 (4.28) | 7.61 (4.97) | 0.00 (0.00) |
| IVAL | 5.31 (0.89) | 5.40 (0.92) | 5.27 (0.87) |
| SEFF | 4.63 (0.99) | 4.67 (1.01) | 4.62 (0.99) |
| TANX | 3.63 (1.36) | 3.60 (1.47) | 3.64 (1.32) |
| CSUS | 4.80 (0.69) | 4.70 (0.78) | 4.84 (0.66) |
| SREL | 4.66 (0.75) | 4.59 (0.84) | 4.69 (0.72) |
| DM | 4.50 (1.00) | 4.41 (1.07) | 4.53 (0.97) |
| DS | 4.50 (0.93) | 4.38 (1.06) | 4.54 (0.88) |
| SM | 3.12 (1.21) | 2.80 (1.26) | 3.25 (1.17) |
| SS | 3.72 (1.08) | 3.58 (1.13) | 3.77 (1.06) |

**Legend:** intrinsic value (IVAL), self-efficacy (SEFF), test anxiety (TANX), cognitive strategy use (CSUS), self-regulation (SREL), deep motive (DM), deep strategy (DS), surface motive (SM), and surface strategy (SS).

**Table 2:** Association between CLAS use and midterm marks

| Coeff | Unstandardized coefficients B ± SE | Standardized coefficients β | t-value | p-value |
|---|---|---|---|---|
| Intercept | 13.58 ± 0.34 | - | 40.03 | **<0.001** |
| CLAS Use | 1.85 ± 0.65 | 0.23 | 2.86 | **<0.005** |

$R^2 = 0.0528$, $F(1, 147) = 8.195$, $p = 0.005$, VIF = 1.00

To probe the factors that may have influenced the association between CLAS use and students' grades, the effects of students' self-regulated learning measured by MSLQ and approaches to learning measured by SPQ were also investigated. A significant improvement in the explained variability of the students' midterm grades was achieved when combining the use of CLAS with the learner characterisations derived from the MSLQ and SPQ scales as shown in Table 3 and Table 4. The two models explain approximately 13% of the variation of the students' midterm grades.

The main contribution was the detection of a statistically significant association between the events in CLAS and the midterm marks when mediated by the categories provided by MSLQ and SPQ (as covariates). This association provides a solid evidence to base the scaffolding of the activities so that a larger percentage of students take advantage of the tool and the activities. Further discussion and practical recommendations are provided in Section 6 - discussion and findings of this report.

**Table 3:** Association between CLAS and midterm marks after controlling for the five MSLQ scales

| Coeff | Unstandardized coefficients B ± SE | Standardized coefficients β | t-value | p-value |
|---|---|---|---|---|
| Intercept | 17.92 ± 2.34 | - | 7.655 | **<0.001** |
| CLAS Use | 1.60 ± 0.64 | 0.20 | 0.6406 | **0.013** |
| IVAL | 0.32 ± 0.46 | 0.08 | 0.696 | 0.488 |
| SEFF | 0.33 ± 0.38 | 0.09 | 0.870 | 0.387 |
| TANX | -0.51 ± 0.22 | -0.19 | -2.287 | **0.024** |
| CSUS | -1.21 ± 0.61 | -0.23 | -1.989 | **0.049** |
| SREL | 0.03 ± 0.54 | 0.01 | 0.059 | 0.953 |

$R^2$ = 0.1341, F(6, 142) =3.664, p = 0.002, VIF =1.154

Legend: Intrinsic value (IVAL), self-efficacy (SEFF), test anxiety (TANX), cognitive strategy use (CSUS), and self-regulation (SREL)

**Table 4:** Association between the use of CLAS and the midterm after controlling for the SPQ scales

| Coeff | Unstandardized coefficients B ± SE | Standardized coefficients β | t-value | p-value |
|---|---|---|---|---|
| Intercept | 17.59 ± 1.77 | | 9.947 | **<0.001** |
| CLAS Use | 1.65 ± 0.64 | 0.21 | 2.570 | **0.011** |
| DM | -0.24 ± 0.40 | -0.07 | -0.595 | 0.553 |
| DS | 0.09 ± 0.42 | 0.02 | 0.217 | 0.829 |
| SM | -0.05 ± 0.38 | -0.02 | -0.133 | 0.894 |
| SS | -0.85 ± 0.42 | -0.25 | -2.028 | **0.044** |

R2 = 0.1287, F(5, 143) = 4.224, p = 0.001, VIF = 1.148

Legend: deep motive (DM), deep strategy (DS), surface motive (SM), and surface strategy (SS).

# 5 Case Study 3: Professional Development Program

Case study 3 draws on a pilot of the CLAS tool for the UNSW Australia's Foundations of University Learning and Teaching Program (FULT) offered in semester 2 of 2014 and semester 1 of 2015. The program was recently redesigned from a traditional on-campus model to be offered in a blended format (combination of face-to-face and online activities). The course redesign models exemplary online and blended learning practice for UNSW academic community. Blended learning approaches apply student-centred learning theories and strategies to enhance student engagement with and understanding of learning material by including collaborative activities, both face-to-face during class time, and asynchronously online outside of class. The program consisted of approximately 60-70 participants each semester comprising of early career academics, sessional teaching staff, HDR students, and post-doctoral fellows. FULT participants were UNSW staff and came from a range of UNSW Faculties and disciplinary areas. Participants nominated themselves for enrolment into this non-accredited program. One of the current strategic goals of the university is to provide blended and flexible learning opportunities for all students. Hence, programs targeted at improving learning and teaching must include opportunities for teaching staff to build capacity in utilizing educational technology to advance students' learning experiences. Further, as many of UNSW's teaching staff do not have formal training in education, the professional development program not only introduces participants to the theoretical concepts of deep learning approaches but must also provide opportunities for participants to engage in activities in a deep and reflective way through blended delivery. One such way, is through the use of a video annotation application, CLAS, which takes a traditionally passive learning experience (watching videos) to a more social and cognitive engaging one (making comments while watching videos) (Project Outcome 1a). The application allows students to make time-stamped annotations and general comments about the video they are watching. These annotations and comments are visible to their peers and the program facilitators empowering both students and faculty to monitor progression (Project Outcome 2b).

## 5.1 Description of how CLAS was integrated into the program design

The FULT program was offered in a blended mode with Moodle as the primary online learning environment. A flipped classroom approach was taken whereby participants were asked to engage with resources and individual activities prior to and subsequently after attending the face-to-face sessions where most of the collaborative activities and discussions occurred. The FULT modules encouraged deep learning through the activities designed and reflective practice with the final task required for completion consisting of an e-portfolio evidencing the participants' progression through the program. FULT consisted of

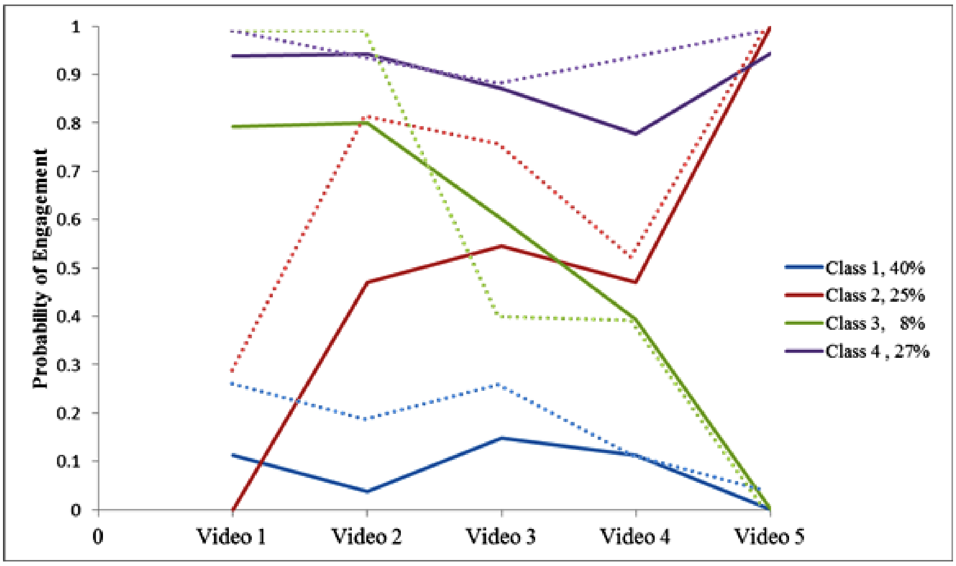
5 modules with one video annotation activity associated with each. The duration of the modules ranged from 2-3 weeks with the face-to-face component in the middle.

To develop academics' skills to use online technologies and make them aware of different online learning strategies to implement in their own blended courses while also reflecting on teaching strategies, the use of a video annotation tool was piloted in the program. Appendix B offers a sample of the detailed instructions provided to students on Moodle. Through the use of a video annotation tool, CLAS, participants were instructed to post time-stamped annotations on the particular segments of the video that they can relate to or consider interesting while watching the videos. They were then asked to review their time-stamped annotations and make a general annotation (non-time stamped) sharing their overall reflection on the video. Participants could view their peers' time and general annotations as well as their own.

## 5.2 Engagement Patterns with CLAS

In this case study, we wanted to explore what usage profiles learners have when interacting with the video annotation software. While 88 participants was involved in the program during Semester 1 in 2014, only 67 participants engaged with the video annotation tool and hence, this case study reports on the engagement patterns of these 67 participants with the video annotation tool for the purposes of investigating the probability of participants engaging with the time-stamped and general annotations. The findings will help inform future iterations of the program. Specifically, for this study, we used two data sources. The *first* one consisted of the digital footprint produced while users interact with CLAS by creating general comments and time-stamped annotations (Guo, Kim, & Rubin, 2014). The *second* data source included responses to the questionnaires that also included items about the participants' approaches to leaning and approaches to teaching. None of these variables produced statistically significant effects in our analyses. These questionnaires were administered at the beginning of both learning experiences reported in this document.

To identify engagement patterns with CLAS through the analysis of the first source of data mentioned above, we made use of latent class analysis (LCA). LCA is a subset of structural equation modelling used to find classes or subtypes of cases in multivariate data (McCutcheon, 1987). This analysis has been widely used to establish a typology of practices/profile/characteristics of individuals or phenomena. A four class solution of LCA was found as a model that provided the best fit with our data. Figure 7 shows the four latent classes that emerged and probability of engagement with the 5 videos. The solid lines represent participants' engagement with general annotations while the dotted lines represent engagement time-stamped annotations. The results show that the majority of participants were in Class 1 (40%).




*Legend: straight lines = engagement with general annotations; dotted lines = engagement with time-stamped annotations*

**Figure 7:** Latent classes of participants across time.

The patterns of engagement with general and time-stamped annotations illustrated in Figure 7 shows differences amongst the 4 classes of participants which can be better interpreted as the following groups or classes of engagement activity:

- Class 1 (minimalists) – Representing the greatest number of participants, this class represents participants who have the highest probability of minimally engaging with both time-stamped and general annotations.
- Class 2 (deferred users) – The second class or group represents participants who tend to have a slower or deferred start with engagement with both time-stamped and general annotations with a particular rise in engagement with the second and fifth videos.
- Class 3 (disenchanted) – This third class representing the smallest proportion of participants in this study, refers to those who have a tendency to engage with the video annotation activity (both time-stamped and general annotations) quite highly at the beginning starting with the first video but who gradually reduce their engagement leading to eventually disengaging entirely by the time they reach the fifth video.
- Class 4 (task-focused) – The fourth and final class of participants represents the participants who have a tendency to be consistently highly engaged with the video annotation activity (both the time-stamped and general annotations) across all five videos. This also means that they have a higher probability of staying on task (e.g. following the instructions provided) and hence can be labelled the *task-focused*.

### 5.2.1 Covariates as predictors of class membership

To explore covariates as predicators of class or group membership discussed in section 5.2, three variables, based on available questionnaire data, were tested to determine their ability to predict class membership: position in the university, age, and gender. Of the three, only participants' position in the university showed to predict their membership in one of the 4 classes: minimalists, deferred users, disenchanted users, and task-focused.

Figure 8 illustrates participants' position in the university and their probability of class membership. It reveals that participants' academic position has the highest probability of influencing class membership for Classes 1 (the minimalists) and 4 (the task-focused) in particular. Specifically casual academics and associate lecturers (generally early career academics) have a tendency to fall in Class 4 while senior lecturers and associate professors tend to fall in class 1. This means that if the professional development program largely consisted of casual academics or associate lecturers, then facilitators could expect a high level of engagement with the video annotation activity. On the other hand, if the program consisted of more senior academic staff, facilitators may need to modify their instructions for the use of the video annotation activity or adapt how they engage participants with the activity in order to increase engagement.

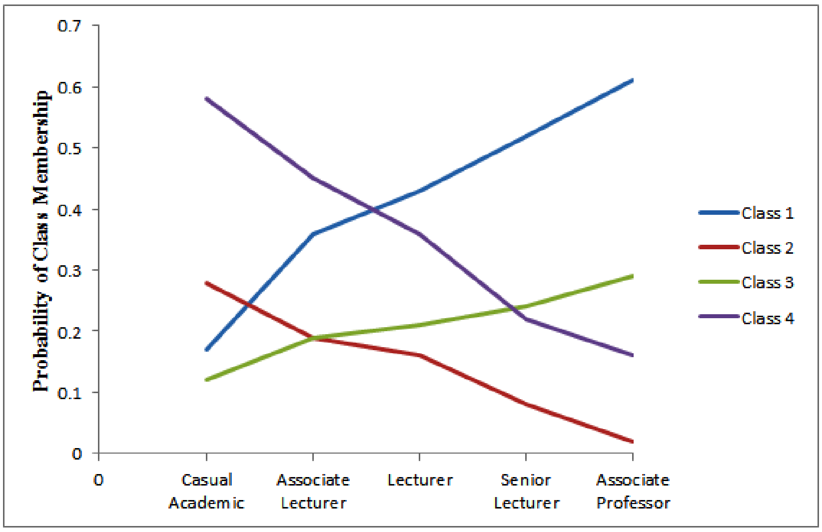



**Figure 8:** Plot of class membership as influenced by the participants' current position at the university

# 6 Discussion and Findings

Several findings and associated recommendations for teaching and learning practice are derived from the three case studies.

*Finding 1: A student's first encounter with a learning technology must be well scaffolded.*

A major recommendation from the first case study relates to the importance of well-designed instructional scaffolds and instructor feedback. The adoption of video annotation is not only driven by student motivation. The study concludes that when students are first exposed to a technology their use must be well scaffolded and integrated into the course design via assessment practices and feedback. This offers some evidence for our proposition that providing a scaffold is necessary (Azevedo & Hadwin, 2005; Beed et al., 1991) for guiding students to appropriately use a tool to aid their learning. This approach reflects Winne. 's (2006) first three conditions for tool adoption – awareness of the tool availability, mapping to a task, and skill to use the tool. Once these conditions are satisfied students will use the tool extensively and appropriately in their learning endeavours. This conclusion is drawn from the significant differences in tool use measured by the number of annotations created and the length of annotation text in Courses 1 and 2. Case study one also noted a significant increase in tool use for the students who transitioned from a course using the tool for grading to subsequent course with the same grading approach. The students in Course 3 – with the continuation of the assessed tool use – almost doubled the number of annotations after their prior experience in Course 2 with the assessed condition. Moreover, the students with that same previous experience from Course 2 maintained the quality of their learning products and increased their level of metacognitive monitoring when enrolled in Course 3. This practice will help sustain higher levels of metacognitive monitoring, which are critical components for learning success and serve to develop student proficiency in self-regulated learning.

*Finding 2: Continuous scaffolding of the video annotation tool use is necessary to assist learners to become effective self-regulated learners.*

In the coding instrument used in the first case study, goals were the most specific type of self-feedback and reflected the highest quality of feedback according to the existing research (Shute, 2008). The findings of the first case study emphasize the importance of continued scaffolding (formative feedback and grading) across several courses in addition to the prior experience with the video annotation tool in a scaffolded condition. This scaffolding is necessary in order to guide students to become more effective at evaluating the products of their own learning (Winne & Hadwin, 1998). The students will continue to increase the accuracy of their judgments of their individual learning progression if they are able to set goals for their future study while self-evaluating their previous learning experience. If student learning goals are not set, students may erroneously stop studying even if they are yet to master the content or procedures under revision.

*Finding 4: Dealing with existing conceptions of learning is important in order for learners to accept new learning strategies with technology such as video annotation.*

Case study two showed that the overall use of video tool software had a positive effect on student grades. The study also showed significant and negative effects of the two covariates on the students' grade – cognitive strategy and test anxiety – two scales of the motivational strategy for learning questionnaire. While the negative effect for test anxiety was anticipated, the effect for cognitive strategy use was not. Students with higher scores of self-reported cognitive strategy use can be seen to be more "hesitant" or reluctant to adopt a new learning tool or approach (e.g. annotating videos). Moreover, we found a significant (yet weak) correlation ($r=-0.161$, $p=0.05$) between the scores of cognitive strategy use and the scores on the midterm exam. These findings suggest that the instructional design needs to offer some further guidance to the students on how to use the new tool effectively and provide a clear rationale outlining the value that such tool use can play in facilitating student learning and therefore, future academic performance. Unlike the first case study with performing arts students, the participatory nature of the second case study lacked the necessary impetus for the first year engineering students with established and effective cognitive strategies to experiment and adopt an alternate process. The findings from this case study further support the results of case study 1.

*Finding 5: Overcoming test anxiety in order to avoid hampering positive effects of the video annotation software.*

Strategies that reduce students' test anxiety are necessary, even in the context of flipped classrooms. Further research examining the association between academic performance, cognitive strategy use, and tool use is needed. It may be the case that there is an indirect effect of cognitive strategy use on academic performance mediated by the tool use. In that mediation, the count of activities (e.g. annotations) may not be the best proxy of learning effort and motivation (as shown in a study by Devolder et al., 2012). These authors demonstrated that the time spent on self-regulated learning within an online learning system mediated the effect of learning experience with the tool on academic performance.

*Finding 6: Promoting effective self-regulated learning with video annotation needs to go hand in hand promoting deep approaches to learning.*

The second case study also examined the effects of the approaches to learning as covariates of the association between the video annotation tool use and students' grade. Only one factor, surface strategy (aside from the use of CLAS), was statistically significant with a negative coefficient. The main conclusion derived from this result is that the students who engaged with CLAS (e.g., made annotations) but retained a surface strategy towards learning had lower midterm scores. These associations suggest that in order to increase midterm results, some measures to promote a deeper approach to learning are also required. Specific promotion of deep approaches to learning is necessary in instructional materials and the overall instructional stance.

*Finding 7: Individual differences play a significant role in the video annotation tool use and drive a need for different instructional strategies to reach out to different student groups.*

The third case study found that four clusters of participants can be used to name the typology of participants' engagement with video annotation across time. Course designers and convenors of academic development programs need to consider that if there are several videos used in the course across the entire teaching and learning period, certain groups of academics have different levels of engagement. Hence, course convenors would not expect that participants who engaged in the beginning of the course with the video annotation activity will sustain their level of engagement until the end of the course.

The covariate related to participants' academic appointment (e.g. casual lecturer through to associate professor) used in this study can be used to predict the most likely cluster membership of the participants. Knowing these characteristics ahead of time would enable course conveners to identify specific interventions to enhance and sustain their engagement in either general or time-stamped annotation or which groups of academics are more likely to engage with such an activity and hence inform future program design.

The findings described in this section inform several recommendations for:

1. The deployment of video annotation tasks must be accompanied by well developed scaffolding (instructional and assessed) to promote student reflection on tool use and learning progress.
2. The adoption of specific measures to convey the purpose and potential advantages of video annotation tasks will faciliate student proficiency in SRL.
3. The inclusion of video annotation tools is a pedagogically sound strategy to develop deep approaches to learning.
4. Video analytics can be applied to provide accurate predictions of student user profiles. These analyses can aid educators in better adapting course design and associated activities for more personalised learning and feedback.

# 7 References


Azevedo, R., & Hadwin, A. F. (2005). Scaffolding Self-regulated Learning and Metacognition – Implications for the Design of Computer-based Scaffolds. Instructional Science, 33(5-6), 367–379. http://doi.org/10.1007/s11251-005-1272-9

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. Retrieved from http://CRAN.R-project.org/package=lme4

Beed, P. L., Hawkins, E. M., & Roller, C. M. (1991). Moving Learners toward Independence: The Power of Scaffolded Instruction. The Reading Teacher, 44(9), 648–655.

Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F. British Journal of Educational Psychology, 71(1), 133–149. http://doi.org/10.1348/000709901158433

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. Annual Review of Psychology, 64(1), 417–444. http://doi.org/10.1146/annurev-psych-113011-143823

Colvin, C., Rogers, T., Wade, A., Dawson, S., Gašević, D., Buckingham Shum, S., … Fisher, J. (2015). Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement (Research Report). Canberra, Australia: Office of Learning and Teaching, Australian Government.

Commonwealth of Australia. (2009). Transforming Australia's higher education system. Barton, ACT: Australian Government.

Conole, G., & Alevizou, P. (2010). A literature review of the use of Web 2.0 tools in Higher Education (A report commissioned by the Higher Education Academy). Milton Keynes, UK: The Open University.

Cummins, J., Brown, K., & Sayers, D. (2007). Literacy, technology, and diversity: Teaching for success in changing times. Boston, MA: Pearson.

Dawson, S., Macfadyen, L., Risko, E. F., Foulsham, T., & Kingstone. (2012). Using technology to encourage self-directed learning: The Collaborative Lecture Annotation System (CLAS). In Proceedings of the Australian Society for Computers in Learning in Tertiary Education Annual Conference (ascilite 2012) Conference (pp. 246–255). Wellington, New Zealand.

de Bruin, A. B., Thiede, K. W., Camp, G., & Redford, J. (2011). Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. Journal of Experimental Child Psychology, 109(3), 294–310.

Devolder, A., van Braak, J., & Tondeur, J. (2012). Supporting self-regulated learning in computer-based learning environments: systematic review of effects of scaffolding in the domain of science education. Journal of Computer Assisted Learning, 28(6), 557–573. http://doi.org/10.1111/j.1365-2729.2011.00476.x

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. Learning and Instruction, 22(4), 271–280.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. Annual Review of Psychology, 53(1), 109-132.

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. Journal of Educational Technology & Society, 15(3), 42–57.

Guo, P., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of mooc videos. In Proceedings of the first ACM conference on Learning at Scale (pp. 41–50). Retrieved from http://dl.acm.org/citation.cfm?id=2566239

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. Metacognition and Learning, 2(2-3), 107–124. http://doi.org/10.1007/s11409-007-9016-7

Hastie, T., Tibshirani, R., & Friedman, J. (2011). The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (2nd ed. 2009. Corr. 7th printing 2013 edition). New York, NY: Springer.

Hulsman, R. L., Harmsen, A. B., & Fabriek, M. (2009). Reflective teaching of medical communication skills with DiViDU: assessing the level of student reflection on recorded consultations with simulated patients. Patient Education and Counseling, 74(2), 142–149. http://doi.org/10.1016/j.pec.2008.10.009

Janssen, J., Erkens, G., & Kirschner, P. A. (2011). Group awareness tools: It's what you do with it that matters. Computers in Human Behavior, 27(3), 1046–1058.

Jovanovic, J., Bagheri, E., & Gasevic, D. (2015). Comprehension and Learning of Social Goals Through Visualization. IEEE Transactions on Human-Machine Systems, 45(4), 478–489. http://doi.org/10.1109/THMS.2015.2419083

Kruger, J., & Dinning, D. (1999). Unskilled and unaware of it: how differences in recognising ones own incompetence lead to inflated self-assessments. Journal of Personality and Social Psychology, 77(6), 1121–1134.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing Pedagogical Action Aligning Learning Analytics With Learning Design. American Behavioral Scientist, 57(10), 1439–1459. http://doi.org/10.1177/0002764213479367

McCutcheon, A. L. (1987). Latent class analysis. Quantitative Applications in the Social Sciences Series No. 64. Thousand Oaks, California: Sage Publications. ISBN 0-521-59451-0.

Mirriahi, N., & Dawson, S. (2013). The pairing of lecture recording data with assessment scores: a method of discovering pedagogical impact. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 180–184). ACM.

Pardo, A., & Kloos, C. D. (2011). Stepping out of the Box: Towards Analytics Outside the Learning Management System. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 163–167). New York, NY, USA: ACM. http://doi.org/10.1145/2090116.2090142

Pinheiro, J., & Bates, D. (2009). *Mixed-Effects Models in S and S-PLUS* (1st ed. 2000. 2nd printing 2009 edition). New York: Springer.

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33–40. http://doi.org/10.1037//0022-0663.82.1.33

Pintrich, P. R., & Zusho, A. (2007). Student motivation and self-regulated learning in the college classroom. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 731–810). Springer.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (Second Edition edition). Thousand Oaks: SAGE Publications, Inc.

Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition*, *28*(6), 1004–1010.

Richardson, J. T. E. (2004). Methodological Issues in Questionnaire-Based Research on Student Learning in Higher Education. *Educational Psychology Review*, *16*(4), 347–358. http://doi.org/10.1007/s10648-004-0004-z

Risko, E. F., Foulsham, T., Dawson, S., & Kingstone, A. (2013). The Collaborative Lecture Annotation System (CLAS): A New TOOL for Distributed Learning. *IEEE Transactions on Learning Technologies*, *6*(1), 4–13. http://doi.org/10.1109/TLT.2012.15

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development*, *52*(1), 5–18. http://doi.org/10.1007/BF02504769

Seltman, H. J. (2012). *Experimental design and analysis*. Online. Retrieved from http://www.stat.cmu.edu/hseltman/309/Book/Book.pdf

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, *78*(1), 153–189. http://doi.org/10.3102/0034654307313795

Siemens, G., & Long, P. D. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, *46*(5), 31–40.

Szpunar, K. K., Jing, H. G., & Schacter, D. L. (2014). Overcoming overconfidence in learning from video-recorded lectures: Implications of interpolated testing for online education. *Journal of Applied Research in Memory and Cognition*, *3*(3), 161–164.

Trigwell, K., Prosser, M., & Waterhouse, F. (1999). Relations between teachers' approaches to teaching and students' approaches to learning. *Higher Education*, *37*(1), 57–70. http://doi.org/10.1023/A:1003548313194

Winne, P. H. (2006). How Software Technologies Can Improve Research on Learning and Bolster School Reform. *Educational Psychologist*, *41*(1), 5–17. http://doi.org/10.1207/s15326985ep4101_3

Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist*, *45*(4), 267–276.

Winne, P. H. (2011). A cognitive and metacognitive analysis of self-regulated learning. In D. H. Schunk & B. Zimmerman (Eds.), *Handbook of self-regulation of learning and performance* (pp. 15–32). New Work and Oxon: Routledge.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

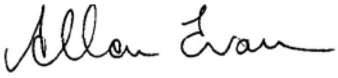
Xu, R. (2003). Measuring explained variation in linear mixed effects models. *Statistics in Medicine*, *22*(22), 3527–3541. http://doi.org/10.1002/sim.1572

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183.

# Appendix A

***Certification by the Provost and Chief Academic officer (or equivalent)***

I certify that all parts of the final report for this OLT grant provide an accurate representation of the implementation, impact and findings of the project, and that the report is of publishable quality.

Signature: 

Name: Professor Allan Evans Date 18 August 2015

# Appendix B

***A sample of the instructions communicated to the studies in the UNSW case study.***

In each module, there will be a required video activity where you will be asked to use a video annotation tool to view a short video, annotate or make time-stamped notes of the concepts introduced that resonate with your teaching or raise questions for you and post a short summary of your annotations or notes to share with your peers. You will be asked to discuss your annotations and comments at the following face-to-face session. The video annotation tool is hosted kindly by the University of South Australia (UniSA), available to UNSW through the Australian Access Federation (AAF), and funded by the Australian Office of Learning and Teaching, please follow the steps below.

Step 1. Login and Access

Click on the link below ([https://clas.unisa.edu.au](https://clas.unisa.edu.au)) to access the video resources for each module through a video annotation tool hosted by UniSA and available to UNSW through the Australian Access Federation (AAF) so you can use your UNSW ZID to access the tool. Once you go to the url, you will see the following image on the screen.

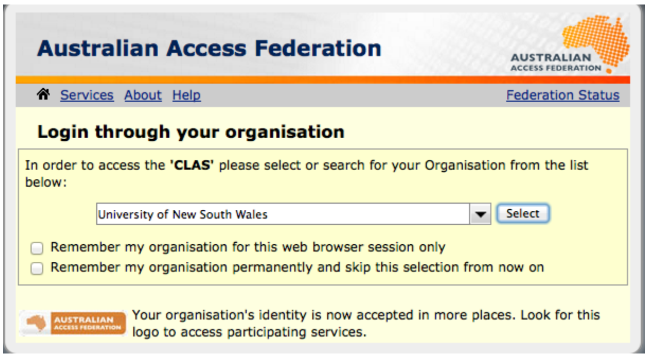



**Figure 9:** AAF Initial login screen

Use the drop down arrow to select The University of New South Wales. You will then be asked to enter your ZID and password to access the tool as shown in this image:

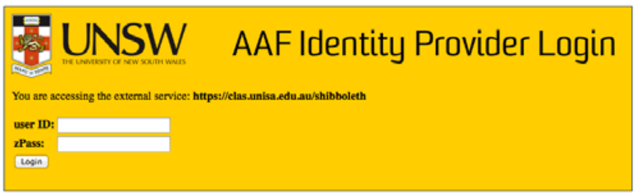



**Figure 10:** Institutional login screen

Step 2: Add annotations.

Once you have successfully logged into the video annotation system, select the <title of video>. Your video will now appear. While watching the video, click on the ‘add annotation’ button to indicate areas (minimum 2-3) of the video that you think introduce a concept that can be applied to your own teaching practice or raise questions for you. (Note: you will be able to view your colleagues' annotations as well). You have the option of providing a 'tag' for your annotation as well - tags can help you search for your annotation later. You can also choose to leave the 'tag' field empty. We encourage you to not mark the annotation as ‘private’ so your peers can also learn from your comments (and you can learn from theirs).

Step 3: Write a General Comment Summary

After finishing the video, review your annotations, and write a brief summary of the concepts discussed in the video that you think can be applied to your own teaching practice or concepts you think cannot be applied (and indicate why or why not). Write this summary in the ‘general comments’ box in the video annotation tool’s interface. We encourage you to not mark this general comment as ‘private’ so your peers can learn from your comments and you can learn from theirs.

The image below shows how the CLAS interface appeared after participants had made time-stamped annotations and general comments with an example of hovering over one of the time-stamped annotation in order to read it:

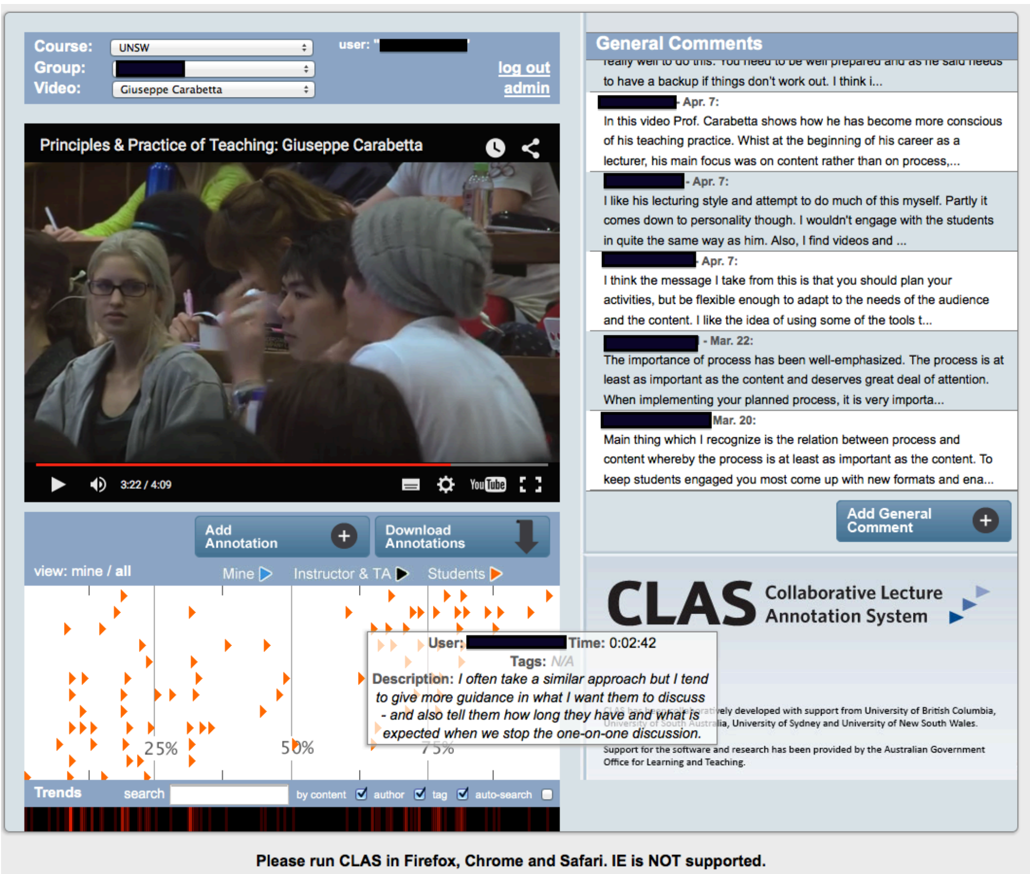



**Figure 11:** User content in the professional development course

# Appendix C

***Statistical analyse of effects of instructional conditions on the adoption and sustained use***

The results reported in Table 5 outline the effects of instructional conditions (assessed vs non-assessed) on the level of the operations performed by students. Although the median value of the number of videos annotated was consistent with the course requirements (three and four videos for Course 1 and Course 2, respectively, i.e., 25% higher requirements in Course 2), the time spent watching videos and the number of operations performed by the students when their general reflective annotations were assessed was by far greater than the increase of 25% in the requirements in Course 2. The effect size of the assessed general video annotations, or the assessed instructional condition (columns A and B) was significant and large for all the operations according to Cohen's interpretation; the only exception was the number of pause events between independent samples of Course 1 and Course 2 (i.e., Course 1a vs. Course 2a). It is interesting to observe that the effect size of the instructional conditions for the students who took both Course 1 and Course 2 in parallel was relatively smaller than between the students who took either Course 1 or Course 2 only. The notable exceptions were only the number of videos annotated: first, the students, who took both Course 1 and Course 2 in parallel, had a tendency to annotate less videos in Course 1 than the students who took Course 1 only (r=.53 vs. r=.75); and second, the effect size on the number of rewind events was almost the same (r=.61 vs. r=.63). However, the comparisons of the students within the same instructional conditions (columns C and D) unveiled no significant differences. Although not significant statistically, the effect size was moderate for time spent watching (r=.32) and the number of rewind events (r=.31) in Course 1 indicating a different pattern in the use of the students who took Course 2 in parallel with Course 1.

Table 6, below, shows the results of the effects of the instructional conditions (assessed vs non-assessed) on the learning products created (i.e., time-stamped annotations and general reflective annotations) and learning strategy followed by the students. A similar trend (column A and B) identified for the operations was also observed for the LIWC variables (WC and Sixltr) with the lower effect size of the assessed instructional condition for the students who took both Course 1 and Course 2 in parallel. However, the students who took both courses at the same time, had a significantly fewer number of words per annotation (WC/Ann) in Course 1 than in Course 2 (r=.63). This is interesting to observe, especially as there was no significant difference between students who were only enrolled in Course 1 or Course 2 only (Course 1a vs. Course 2a, column A).

**Table 5.** Comparison of the operations between Course 1 and Course 2

| Variables | Course 1a (n=23) | Course 1b (n=8) | Course 2a (n=32) | Course 2b (n=8) | A<br>B | C<br>D |
|---|---|---|---|---|---|---|
| Videos annotated | 3.00<br>(3.00) | 3.00<br>(2.25, 3.00) | 4.00<br>(4.00, 4.00) | 4.00<br>(4.00, 4.00) | U=680.00, z=5.79, p<.001, r=.53<br>T=32.50, z=2.126, p=.033, r=.75 | U=92.00, z=.00, p=1.00, r=.00<br>U=133.00, z=.24, p=.809, r=.04 |
| Annotations | 11.00<br>(8.00, 21.00) | 14.00<br>(8.25, 16.50) | 55.00<br>(47.25, 59.75) | 58.00<br>(53.00, 62.25) | U=717.5, z=5.97, p<.001, r= .81<br>T=36.00, z=2.52 p=.012, r=.63 | U=99.50, z=.34, p=.734, r=.06<br>U=146.50, z=.626, p=.531, r=.10 |
| Time watched (sec) | 1490.00<br>(888.00, 2200.00) | 2262.00<br>(1374.25, 2535.75) | 5135.00<br>(3604.25, 8077.50) | 7338.00<br>(4039.50, 10601.25) | U=687.00, z=5.44, p<.001, r=.73<br>T=36.00, z=2.52, p=.012, r=.63 | U=130.00, z=1.72, p=.086, r=.32<br>U=157.00, z=.98, p=.327, r=.16 |
| Pause | 14.00<br>(7.00, 31.00) | 34.50<br>(17.50, 51.50) | 97.00<br>(59.75, 145.50) | 131.00<br>(37.50, 186.00) | U=554.50, z=3.32, p<.001, r=.34<br>T=36.00, z=2.52, p=.012, r=.63 | U=126.00, z=1.54, p=.125, r=.29<br>U=137.00, z=.30, p=.761, r=.05 |
| Rewind | 9.00<br>(2.00, 25.00) | 27.00<br>(8.25, 38.50) | 43.00<br>(20.25, 78.00) | 78.00<br>(24.00,134.25) | U=633.00, z=4.52, p<.001, r=.61<br>T=36.00, z=2.52, p=.012, r=.63 | U=128.50, z=1.65, p=.099, r=.31<br>U=158.00, z=1.01, p=.310, r=.16 |
| Fast forward | 3.00<br>(1.00, 5.00) | 4.00<br>(.50, 8.50) | 29.00<br>(18.00, 41.00) | 15.50<br>(7.50, 50.50) | U=720.0, z=6.02, p<.001, r=.81<br>T=28.00, z=2.37, p=.018, r=.59 | U=110.50, z=.85, p=.398, r=.16<br>U=105.50, z=-.76, p=.446, r=-.12 |

Legend: A – Course 1a vs Course 2a; B – Course 1b vs Course 2b; C – Course 1a vs Course 1b; D – Course 2a vs Course 2b.

**Table 6.** Comparison of the learning products and learning strategy between Course 1 and Course 2

| Variables | Course 1a (n=23) | Course 1b (n=8) | Course 2a (n=32) | Course 2b (n=8) | A<br>B | C<br>D |
|---|---|---|---|---|---|---|
| WC | 322.00 (108.00, 653.00) | 241.00 (178.75, 280.50) | 1671.50 (1218.50, 2212.25) | 1593.00 (1502.00, 1921.50) | U=721.0, z=6.02, p<.001, r=.81<br>T=36.00, z=2.52, p=.012, r=.63 | U=66.50, z=-1.15, p=.250, r=-.21<br>U=124.00, z=-.12, p=.906, r=-.02 |
| WC/Ann | 23.56 (13.48, 46.91) | 16.65 (14.74, 20.17) | 30.94 (25.42, 39.79) | 29.30 (24.50, 33.13) | U=463.0, z=1.62, p=.105, r=.22<br>T=36.00, z=2.52, p=.012, r=.63 | U=72.00, z=-.90, p=.367, r=-.17<br>U=102.00, z=-.88, p=.379, r=-.14 |
| Sixltr | 175.61 (126.09, 334.89) | 190.37 (97.73, 271.33) | 892.25 (691.57, 044.11) | 878.73 (814.26, 972.09) | U=708.0, z=5.80, p<.001, r=.78<br>T=36.00, z=2.52, p=.012, r=.63 | U=89.00, z=-.14, p=.892, r=-.03<br>U=118.00, z=-.34, p=.735, r=-.05 |
| Sixltr/Ann | 16.57 (14.19, 20.22) | 17.74 (11.73, 19.99) | 15.35 (14.61, 18.49) | 15.36 (14.59, 20.51) | U=321.0, z=-0.80, p=.423, r=--.11<br>T=21.00, z=.42, p=.674, r=.11 | U=91.00, z=-.05, p=.964, r=-.01<br>U=133.00, z=.169, p=866, r=.03 |
| Narrativity | -.42 (-.97, -.04) | -.62 (-.87, -.19) | .27 (-.62, .59) | -.13 (-.92, .14) | U=521.00, z=2.61, p=.009, r=.35<br>T=19.00, z=.14, p=.889, r=.04 | U=85.00, z=-.32, p=.752, r=-.06<br>U=86.00, z=-1.42, p=.156, r=-.22 |
| Deep Cohesion | -1.34 (-1.72, -.79) | -1.55 (-1.83, -.78) | -.60 (-1.01, -.18) | -.58 (-1.15, .11) | U=534.00, z=2.83, p=.005, r=.38<br>T=36.00, z=2.52, p=.012, r=.63 | U=73.00, z=-.86, p=.391, r=-.15<br>U=152.00, z=.81, p=.417, r=.13 |
| Referential Cohesion | .43 (-.11, .60) | .07 (-.31, .69) | .18 (.01, .45) | .44 (.02, .60) | U=303.00, z=-1.11, p=.269, r=-.15<br>T=27.00, z=1.26, p=.208, r=.32 | U=80.00, z=-.54, p=.588, r=-.10<br>U=144.00, z=.54, p=.589, r=.09 |
| Syntactic Simplicity | -.18 (-.48, .12) | .32 (-.16, 1.27) | -.32 (-.48, .14) | -.20 (-.59, .05) | U=337.00, z=-.53, p=.597, r=-.07<br>T=6.00, z=-1.68, p=.093, r=-.42 | U=128.00, z=1.63, p=.104, r=.29<br>U=113.00, z=-.51, p=.612, r=-.08 |
| Word Concreteness | 1.10 (.87, 1.73) | .95 (.39, 1.42) | .42 (-.05, 1.12) | 1.37 (.52, 1.64) | U=521.00, z=-2.61, p=.009, r=-.35<br>T=18.00, z=.00, p=1.000, r=.00 | U=72.00, z=-.90, p=.367, r=-.16<br>U=165.00, z=1.25, p=.211, r=.20 |
| Density | .17 (.13, .20) | .18 (.15, .22) | .25 (.21, .28) | .26 (.18, .31) | U=603.00, z=4.01, p<.001, r=.54<br>T=26.00, z=1.12, p=.263, r=.28 | U=115.00, z=1.06, p=.289, r=.20<br>U=139.00, z=.37, p=.710, r=.06 |

Legend: A – Course 1a vs Course 2a; B – Course 1b vs Course 2b; C – Course 1a vs Course 1b; D – Course 2a vs Course 2b.

The results of the analysis of the Coh-Metrix principal components showed some significant differences between Courses 1 and 2. Moderate effect sizes and significantly higher scores of narrativity (r =.35) and deep cohesion (r=.38) and significantly lower scores for syntactic simplicity (r=-.42) and word concreteness (r=.-35) of the students who took Course 1 only than those of the students who took Course 2 only (Course 1a vs Course 2b) were found (Column A – Table 6). Significant differences within the Coh-Metrix scores of the students who took both Course 1 and Course 2 in parallel were found (Column B – Table 6).

Specifically, significantly higher scores of deep cohesion (r=.63 – large effect size) and referential cohesion (r=.32 – moderate effect size) and significantly lower scores of syntactic simplicity (r=-.42) and word concreteness (r=-.35) for the students in Course 2 compared to their scores in Course 1. No significant differences were found between the scores in Course 1 for the students who took Course 1 only compared to those who took both Course 1 and Course 2 in parallel (Column C – Table 6). Similar non-significant findings apply to the students who took Course 2 only compared to those who took both Course 1 and Course 2 in parallel (Column D – Table 6).

Finally, the learning strategy followed by the students in the assessed condition exhibited a significantly higher level of metacognitive monitoring measured by the transition graph density than that of the students who were in the non-assessed condition only. The effect size of this difference was large (Course 1a vs. Course 2a, r=.54). The significant difference was not unveiled between the density values of the students who took Course 1 and Course 2 in parallel, although the effect size was marginally moderate (r=.28).

Table 7, below, shows the effects of students' prior experience with a learning tool, CLAS on the operations they perform when they use the tool in a subsequent course with the same assessed instructional conditions (Course 2 vs. Course 3).

**Table 7.** Comparison of the operations between Course 2 and Course 3

| Variables | Course 2c (n=22) | Course 2d (n=18) | Course 3a (n=10) | Course 3b (n=18) | A<br>B | C<br>D |
|---|---|---|---|---|---|---|
| Videos annotated | 4.00 (3.75, 4.00) | 4.00 (4.00, 4.00) | 5.00 (4.00, 5.00) | 5.00 (4.00, 5.00) | U=194.50, z=3.98, p<.001, r=.70<br>T=70.00, z=1.941, p=.052, r=.32 | U=268.50, z=2.75, p=.006, r=.43<br>U=80.50, z=-.54, p=.587, r=-.10 |
| Annotations | 55.00 (40.25, 58.50) | 58.00 (49.75, 63.25) | 92.00 (80.25, 95.25) | 104.50 (88.25, 109.00) | U=210.50, z=4.09, p<.001, r=0.72<br>T=171.00, z=3.72, p<.001, r=.62 | U=247.50, z=1.35, p=.178, r=.21<br>U=128.00, z=1.85, p=.065, r=.35 |
| Time watched (sec) | 4887.50 (2873.50, 8139.50) | 6136.00 (3984.75, 8427.75) | 5395.50 (5119.25, 8753.50) | 7124.50 (3831.75, 9256.50) | U=113.00, z=.33, p=.345, r=.06<br>T=101.00, z=.68, p=.500, r=.11 | U=233.00, z=.95, p=.341, r=.15<br>U=98.00, z=.38, p=.701, r=.07 |
| Pause | 86.50 (55.75, 157.00) | 103.00 (78.75, 139.50) | 112.50 (77.25, 153.50) | 123.50 (89.25, 155.25) | U=120.50, z=.43, p=.675, r=.08<br>T=123.50, z=1.656, p=.098, r=.28 | U=213.50, z=.42, p=.673, r=.07<br>U=99.00, z=.43, p=.666, r=.08 |
| Rewind | 43.00 (24.25, 81.25) | 46.50 (20.75, 109.00) | 54.00 (31.50, 110.50) | 49.00 (26.00, 97.50) | U=124.50, z=.59, p=.562, r=.10<br>T=96.50, z=.96, p=.344, r=.16 | U=208.00, z=.27, p=.786, r=.04<br>U=87.00, z=-.144, p=.886, r=-.03 |
| Fast forward | 28.50 (9.75, 48.00) | 25.50 (10.75, 42.25) | 19.00 (15.00, 35.00 | 29.50 (15.25, 37.25) | U=92.00, z=.73, p=.489, r=.13<br>T=81.50, z=.24, p=.813, r=.04 | U=177.00, z=-.57, p=.568, r=-.09<br>U=97.00, z=.336, p=.737, r=.6 |

Legend: A – Course 2c vs Course 3a; B – Course 2d vs Course 3b; C – Course 2c vs Course 2d; D – Course 3a vs Course 3b.

Specifically, the table focuses on the operations performed and time spent. The only significant effects (Column A and B of Table 7) were observed for the number of videos annotated and the total number of annotations created. While the increase in the number of videos annotated was expected due to the increase in total videos in Course 3 (five videos) compared to Course 2 (four videos), the number of annotations created was much higher (almost doubled) than the increase in each student's performance video and hence, the course requirements for the tool use. Interestingly, the effect size of the instructional condition in Course 3 was larger for the students who had previously been enrolled in Course 2 which had the same instructional condition, assessed general annotations (r=.72 vs. r=.62). Moreover, although no significant differences were observed between the students in the same condition (Column C and D), the moderate effect size of the previous enrolment in Course 2 (i.e., previous experience using CLAS for assessed general annotations) on the number of annotations created by students in Course 3 is notable (r=.35, Table 4 – Colum D). That is, students who used the tool for the second time (Course 3b) in the same assessed instructional condition had a tendency to create more annotations than the students who used the tool in the assessed instructional condition for the first time (Course 2d).

Table 8 presents the results of the effects of students' prior use of a learning tool, CLAS, on their learning products created and learning strategy followed when they used the tool in subsequent courses with the same assessed instructional condition (Course 2 vs. Course 3). As expected with the increase of the annotations created in Course 3, the number of words and words longer than six letters was greater in Course 3 compared to Course 2. Interestingly, the effect size for the students who previously used the tool in Course 2 was higher than for those who used the tool for the same time in Course 3. This is further reflected by the significant difference with the moderate effect size (r=.35) between the number of words per annotation for the students who took both Course 2 and Course 3 (Course 2d vs. Course 3b) whose median value increased from Mdn = 31.30 in Course 2 to Mdn=35.75 in Course 3. The comparison of the students within the same condition (Columns C and D in Table 8) showed that students who previously had enrolled only in Course 2 had a significantly greater number of words (WC) in their annotations than those who only enrolled in Course 3 (Course 3a vs. Course 3b) with a moderate effect size (r=.41).

The results of the analysis of the Coh-Metrix principal components showed some significant differences between Courses 2 and 3. Significantly lower scores of narrativity (r=-.40) and deep cohesion (r=-.47) and significantly higher scores of concreteness (r=.32) for the students who took Course 3 only compared to those who Course 2 only (Course 2c vs Course 3a) were found (Column A – Table 8). No significantly different scores between the Course 2 and Course 3 scores of students who took both Course 2 and Course 3 were detected (Column B – Table 8). Likewise, no significant differences in the Coh-Metrix scores in Course 2 between the students who took Course 2 only versus those who took both Course 2 and 3 were found (Column C – Table 8). Significantly higher scores for referential cohesion (r=.43) with moderate effect size in Course 3 for the students who took Course 3 only compared to the Course 3 scores of the students who took both Course 2 and 3 were observed.

**Table 8.** Comparison of the learning products and strategy between Course 2 and Course 3

| Variables | Course 2c (n=22) | Course 2d (n=18) | Course 3a (n=10) | Course 3b (n=18) | A<br>B | C<br>D |
|---|---|---|---|---|---|---|
| WC | 1452.00 (1071.50, 2196.75) | 1829.50 (1551.50, 2101.25) | 2604.50 (2106.00, 3267.75) | 3224.50 (2961.75, 4230.75) | U=182.00, z=2.85, p=.003, r=.50<br>T=171.00, z=3.72, p<.001, r=.62 | U=264.50, z=1.81, p=.071, r=.29<br>U=135.00, z=2.16, p=.031, r=.41 |
| WC/Ann | 28.31 (24.92, 38.84) | 31.30 (27.45, 36.70) | 29.74 (22.70, 37.14) | 35.75 (27.46, 40.64) | U=101.00, z=-0.37, p=.714, r=-.07<br>T=134.00, z=2.11, p=.035, r=.35 | U=216.00, z=.49, p=.625, r=.08<br>U=123.00, z=1.58, p=.114, r=.30 |
| Sixltr | 835.62 (600.20, 1046.19) | 911.56 (830.68, 1030.83) | 1492.03 (1035.95, 1597.58) | 1549.82 (1277.99, 1791.60) | U=192.00, z=3.33, p=.001, r=.59<br>T=171.00, z=3.72, p<.001, r=.62 | U=247.00, z=1.33, p=.183, r=.21<br>U=195.00, z=1.30, p=.195, r=.24 |
| Sixltr/Ann | 15.40 (14.91, 18.47) | 15.31 (14.53, 18.75) | 15.87 (11.89, 17.34) | 15.12 (13.89, 16.70) | U=93.00, z=-0.69, p=.489, r=-.12<br>T=49.00, z=-1.59, p=.112, r=-.27 | U=188.00, z=-.27, p=.786, r=-.04<br>U=84.00, z=-.288, p=.774, r=-.05 |
| Narrativity | .26 (-.54, .60) | -.11 (-.85, .46) | -.61 (-.94, .13) | -.20 (-.71, .37) | U=54.00, z=-2.28, p=.023, r=-.40<br>T=89.00, z=.15, p=.879, r=.03 | U=155.00, z=-1.17, p=.242, r=-.18<br>U=114.00, z=1.15, p=.250, r=.22 |
| Deep Cohesion | -.60 (-.87, -.05) | -.61 (-1.32, -.28) | -1.27 (-1.80, -.64) | -1.23 (-1.54, -.32) | U=44.00, z=-2.68, p=.007, r=-.47<br>T=52.00, z=-1.46, p=.145, r=-.24 | U=167.00, z=-.843, p=.399, r=-.13<br>U=111.00, z=1.01, p=.314, r=.19 |
| Referential Cohesion | .18 (-.01, .41) | .34 (.07, .54) | .21 (-.03, .40) | .39 (.30, .55) | U=104.00, z=-.24, p=.807, r=-.04<br>T=106.00, z=.893, p=.372, r=.15 | U=221.00, z=.63, p=.532, r=.10<br>U=137.00, z=2.25, p=.024, r=.43 |
| Syntactic Simplicity | -.27 (-.48, .09) | -.36 (-.54, .11) | .00 (-.51, .41) | -.39 (-.79, .00) | U=131.00, z=.85, p=.393, r=.15<br>T=62.00, z=-1.02, p=.306, r=-.17 | U=180.00, z=-.49, p=.625, r=-.08<br>U=57.00, z=-1.58, p=.114, r=-.30 |
| Word Concreteness | .33 (-.11, 1.03) | 1.02 (.28, 1.69) | 1.10 (.27, 1.63) | 1.28 (.38, 1.42) | U=155.00, z=1.83, p=.067, r=.32<br>T=89.00, z=.15, p=.879, r=.03 | U=260.00, z=1.69, p=.092, r=.27<br>U=88.00, z=-.10, p=.944, r=-.02 |
| Density | .25 (.19, .30) | .25 (.21, .28) | .26 (.25, .32) | .29 (.26, .33) | U=149.50, z=1.61, p=.108, r=.28<br>T=.135, z=2.16, p=.031, r=.36 | U=212.50, z=.39, p=.693, r=.06<br>U=107.50, z=.84, p=.401, r=.16 |

Legend: A – Course 2c vs Course 3a; B – Course 2d vs Course 3b; C – Course 2c vs Course 2d; D – Course 3a vs Course 3b

The analysis did not reveal any significant differences in the level of metacognitive monitoring between the students in Course 2 and Course 3 measured by the density of their transition graphs (Table 8).

The results shown in Table 9, below, show the effects of students' previous assessed experience with a learning tool, CLAS, on the operations they perform in a subsequent course when there is a non-assessed instructional condition (i.e. general reflective annotations are not assessed, Course 2 vs. Course 4). There were no observed differences in the number of videos annotated. This was expected as the number of videos annotated by each student was consistent across the two courses (i.e. four in both of them). Likewise, there was no significant difference in the number of annotations created in the two courses. The comparison between students within Course 4 found a moderate effect (r=.35), although statistically insignificant (p=.138), with those who previously had used CLAS in Course 2 (assessed instructional condition) making a greater number of annotations than those who used the tool in Course 4 for the first time. The numbers of pause, rewind, and fast forward events significantly declined with large effect sizes in Course 4 compared to those of Course 2.

**Table 9.** Comparison of the operations between Course 2 and Course 4

| Variables | Course 2e (n=29) | Course 2f (n=11) | Course 4a (n=9) | Course 4b (n=11) | A<br>B | C<br>D |
|---|---|---|---|---|---|---|
| Videos annotated | 4.00<br>(4.00, 4.00) | 4.00<br>(4.00, 4.00) | 4.00<br>(3.00, 4.50) | 4.00<br>(3.00, 5.00) | U=115.00, z=-.65, p=.613, r=-.11<br>T=18.00, z=.00, p=1.000, r=.00 | U=171.00, z=.50, p=.618, r=.08<br>U=55.50, z=.47, p=.636, r=.11 |
| Annotations | 55.00<br>(48.00, 59.00) | 58.00<br>(43.00, 67.00) | 39.00<br>(18.00, 55.50) | 55.00<br>(40.25, 58.50) | U=86.50, z=-1.51, p=.133, r=-.25<br>T=36.50, z=.31, p=.756, r=.07 | U=185.50, z=.78, p=.430, r=.12<br>U=69.00, z=1.48, p=.138, r=.35 |
| Time watched (sec) | 4878.00<br>(3635.50, 8289.00) | 5863.00<br>(3704.00, 8415.00) | 680.00<br>(236.50, 1163.00) | 4887.50<br>(2873.50, 8139.50) | U=5.00, z=-4.31, p<.001, r=-.70<br>T=5.00, z=-2.49, p=.013, r=-53 | U=167.00, z=.23, p=.820, r=.04<br>U=51.00, z=.115, p=.908, r=.03 |
| Pause | 101.00<br>(53.50, 152.50) | 103.00<br>(57.00, 144.00) | 19.00<br>(3.00, 27.50) | 86.50<br>(55.75, 157.00) | U=15.0, z=-3.97, p<.001, r=-.62<br>T=.00, z=-2.93, p=.003, r=-.63 | U=177.50, z=.55, p=.586, r=.09<br>U=42.50, z=-.54, p=.590, r=-.13 |
| Rewind | 40.00<br>(20.50, 107.50) | 56.00<br>(25.00, 79.00) | 3.00<br>(1.50, 11.50) | 43.00<br>(24.25, 81.25) | U=19.00, z=-3.83, p<.001, r=-.68<br>T=.00, z=-2.93, p=.003, r=-.63 | U=172.50, z=.39, p=.694, r=.06<br>U=39.50, z=-.78, p=.436, r=-.18 |
| Fast forward | 29.00<br>(10.00, 43.50) | 26.00<br>(18.00, 48.00) | 5.00<br>(.00, 9.50) | 28.50<br>(9.75, 48.00) | U=23.50, z=-3.68, p<.001, r=-.60<br>T=.00, z=-2.94, p=.003, r=-.63 | U=172.50, z=.39, p=.693, r=.06<br>U=46.00, z=-.27, p=.785, r=-.06 |

Legend: A – Course 2e vs Course 4a; B – Course 2f vs Course 4b; C – Course 2e vs Course 2f; D – Course 4a vs Course 4b

Table 10 outlines the results of the effects of students' previous assessed experience with a learning tool, CLAS, on the learning products created and learning strategy followed in a subsequent course with a non-assessed instructional condition (Course 2 vs. Course 4). No significant decline in the number of words in annotations (WC) and number of words per annotation (WC/Ann) in Course 4 ware found for the students who had prior experience with using CLAS with assessed instructional conditions (Column B of Table 10 – Course 2f vs

Course 4b). However, significant differences were observed between the number of words (WC) and words per annotation (WC/Ann) of students who enrolled in only Course 2 or Course 4 (Course 2e vs. Course 4a). The difference in the number of words per annotation was significant with a large effect size for students in Course 4 with or without previous experience with using CLAS in an assessed instructional condition (Column D, Table 10).

No significant differences of the Coh-Metrix scores were found for the study comparison groups of the students in Courses 2 and 4 as shown in Table 10. Finally, the metacognitive monitoring measured by the density of the transition graphs of students' learning strategy significantly dropped in Course 4 compared to Course 2 with a large effect size, regardless of previous experience with the tool.

**Table 10.** Comparison of the learning products and operations between Course 2 and Course 4

| Variables | Course 2e (n=29) | Course 2f (n=11) | Course 4a (n=9) | Course 4b (n=11) | A<br>B | C<br>D |
|---|---|---|---|---|---|---|
| WC | 1572.00 (1332.50, 2023.50) | 1953.00 (1133.00, 2377.00) | 837.00 (357.00, 1230.50) | 2050.00 (727.00, 3747.00) | U=48.00, z=-2.83, p=.005, r=-.46<br>T=32.00, z=-.09, p=.929, r=-.02 | U=186.00, z=.80, p=.422, r=.13<br>U=73.00, z=1.79, p=.074, r=.42 |
| WC/Ann | 28.51 (25.66, 37.92) | 33.27 (25.12, 38.28) | 21.46 (10.45, 26.32) | 33.14 (21.49, 40.20) | U=33.00, z=-3.35, p=.001, r=-.54<br>T=24.00, z=-.80, p=.424, r=.17 | U=185.00, z=.77, p=.440, r=.12<br>U=81.00, z=2.39, p=.017, r=.56 |
| Sixltr | 873.67 (742.14, 984.30) | 1025.97 (648.81, 1064.02) | 756.28 (294.44, 943.53) | 1033.98 (595.03, 1478.45) | U=86.00, z=-1.53, p=.127, r=-.25<br>T=41.00, z=.711, p=.477, r=.15 | U=195.00, z=1.08, p=.282, r=.17<br>U=68.00, z=1.41, p=.160, r=.33 |
| Sixltr/Ann | 15.38 (14.61, 18.16) | 15.12 (14.45, 18.83) | 15.78 (11.88, 19.87) | 16.95 (16.05, 18.86) | U=123.00, z=-.26, p=.127, r=-.04<br>T=51.00, z=1.60, p=.110, r=.34 | U=159.00, z=-.02, p=.988, r=.00<br>U=58.00, z=.65, p=.518, r=.15 |
| Narrativity | .06 (-.64, .56) | .28 (-.94, .67) | -.32 (-.89, -.01) | -.15 (-.21, -.09) | U=79.00, z=-1.79, p=.077, r=-.29<br>T=30.00, z=-.27, p=.790, r=-.06 | U=169.00, z=.29, p=.774, r=.05<br>U=67.00, z=1.33, p=.184, r=.30 |
| Deep Cohesion | -.64 (-1.10, -.07) | -.49 (-.93, -.33) | -.86 (-1.89, -.17) | -.34 (-.58, -.13) | U=105.00, z=-.88, p=.381, r=-.14<br>T=47.00, z=1.25, p=.213, r=.27 | U=169.00, z=.29, p=.774, r=.05<br>U=68.00, z=1.41, p=.160, r=.31 |
| Referential Cohesion | .29 (-.01, .55) | .17 (.11, .36) | .12 (-.05, .50) | .07 (-.02, .12) | U=109.00, z=-.74, p=.460, r=-.12<br>T=18.00, z=-1.33, p=.182, r=-.28 | U=148.00, z=-.35, p=.728, r=-.06<br>U=43.00, z=-.494, p=.621, r=-.11 |
| Syntactic Simplicity | -.32 (-.54, .12) | -.27 (-.45, .03) | .02 (-.14, .47) | -.04 (-.12, .05) | U=179.00, z=1.67, p=.096, r=.27<br>T=42.00, z=.80, p=.424, r=.17 | U=178.00, z=.56, p=.575, r=.09<br>U=38.00, z=-.87, p=.382, r=-.20 |
| Word Concreteness | .75 (-.15, 1.51)) | .42 (.23, 1.36) | .70 (.09, 1.88) | .32 (.12, .90) | U=156.00, z=.88, p=.381, r=.14<br>T=24.00, z=-.80, p=.424, r=-.17 | U=184.00, z=.74, p=.458, r=.12<br>U=38.00, z=-.87, p=.382, r=-.20 |
| Density | .25 (.21, .28) | .25 (.20, .30) | .06 (.01, .09) | .06 (.05, .11) | U=45.00, z=-4.48, p<.001, r=-.73<br>T=.00, z=-2.93, p=.003, r=-.63 | U=165.50, z=.18, p=.856, r=.03<br>U=64.00, z=1.10, p=.270, r=.26 |

Legend: A – Course 2e vs Course 4a; B – Course 2f vs Course 4b; C – Course 2e vs Course 2f; D – Course 4a vs Course 4b

# Appendix D

***Analysis of Reflection Specificity in Video Annotations***

In the process of content analysis, we realized that the coding instructions provided in the paper by Hulsman et al. (2009) were insufficient for the coders to achieve high validity, and thus the instructions were further clarified to take the form as they are shown in Table 11. The unit of analysis was a part of a sentence. For example, the following sentence had two specificity indicators as marked in the text below:

"*Staying with the ensemble through preps and through releases are very important [goal] and I am not doing those any where near enough [observation].*"

Multi-level linear mixed models are applied for the analysis of the effects of *instructional conditions* and *experience* (as independent variables) on the level of specificity of the students' reflections (i.e., self-evaluation) in the case study with the performing arts students. Mixed-effects modeling provide a robust and flexible approach that allows for a wide set of correlation patterns to be modeled and is recommended method for studying similar datasets (Pinheiro & Bates, 2009; Seltman, 2012). Mixed-effects models include a combination of fixed and random effects and can be used to assess the influence of the fixed effects on dependent variables after accounting for any extraneous random effects. Fixed effects correspond to the numerical or categorical variables that are of primary interest and represent fixed, repeatable levels among which comparisons are to be made. Random effects are categorical variables that represent variability among subjects, a random selection from a larger population to which the results can be extended.

Relying on mixed-effects modeling approach we were able to examine the association between the two conditional factors - instructional design and previous experience – and the two different outcomes – count of each level of student reflection and the total count of general annotations (regardless the reflective state). Thus, for each of the three reflective states (goal, observation, and motive/effect) and the total count of general annotations, we constructed two mixed-effects models. In order to assess whether independent variables predict the activity (count of reflections) above and beyond the random effects, we build two models for each of the dependent variables – *null model* and *final model* (Table 12). The initial, *null model* included the random effect only (*student within a course*). However, only in case when we analyzed the association between count of motive/effect reflections and the independent variables, we were not able to fit the model with such a structure of random effects. Therefore, only in that single case, we specified *student* as a random effect. On the other hand, a *final model* included instructional condition, experience, and interaction between instructional condition and experience, as fixed effects. Intra-class correlation coefficient (ICC) (Raudenbush & Bryk, 2002), second-order Akaike information criterion (AICc) and likelihood ratio test (Hastie, Tibshirani, & Friedman, 2011) were used to decide on the best fitting model (Table 12). We also estimated an effect size ($R^2$) for each model as goodness-of-fit measure, calculating the variance explained using the method suggested by Xu (2003).

**Table 11. Coding scheme used for content analysis of the specificity of reflections**

| Category | Definition | Example |
|---|---|---|
| Observation | The student indicates what they have observed or how they felt or feel about something. This observation may apply to their own behaviour or the behaviour of the other students. They do **not** indicate why the behaviour occurred. | "*I feel like that in this lab I felt more comfortable than before standing in front of this group than in previous lab.*"<br><br>"*I still continue to have problems with making eye contact...*" |
| Motives/Effects | The student indicates what they have observed but also explains **why** it occurred (in the past). It does **not** apply to when a student writes a goal and provides a reason for the goal. | "*Whenever I'm getting tense when I play trumpet I try and simplify what I'm doing*."<br><br>"…, but being up there made me feel insecure and nervous, which led to my eyes dropping more frequently." |
| Goals | The student indicates what they will do next time or what they need to work on (in the future or from now on). This also includes if a student indicates a standard behaviour that should be achieved (in the future). | "*I think for my next lab my goals shall be to try to make eye contact with everyone at least once*."<br><br>"*What I really want to avoid is ending up just mirroring everything.*" |

Linear mixed-effects models were conducted using R v.3.0.1 software for statistical analysis with package lme4 (Bates, Maechler, Bolker, & Walker, 2015). The hypotheses specify the direction of the effect, however two-tailed tests were used for significance testing with an alpha level of .05.

The likelihood ratio test for ***General annotation count*** models yielded significantly better fit of the *final* model (i.e., the model that included fixed and random effects) than the *null* model. The model (Figure 12d) showed a significant effect of the previous experience on the total number of general annotations produced – $F(1, 25.150) = 8.723$, $p = .007$, while the effect of the instructional condition was marginally significant – $F(1, 2.048) = 12.214$, $p = .071$. Observing these results, it could be concluded that the students with previous experience with the tool use tend to produce more annotations in general, than those students who encounter the tool for the first time. This was further confirmed with the pairwise comparison of least square means test that showed significant difference between the two levels of experience: no experience vs. with experience – $t(25.10) = -2.95$, $p = .007$, 95% CI [-8.01, -1.43].

**Table 12. Inferential statistics for the model fit assessment of the effects of instructional condition and experience on the specificity of reflective activities with video annotation**

| Model | $\chi^2$ | df | $R^2$ marg. | $R^2$ cond. | AICc | ICC student | ICC course |
|---|---|---|---|---|---|---|---|
| | | | | ***Goal count*** | | | |
| *null* | | | - | .684 | 832.911 | .104 | .580 |
| ***final*** | **42.866*****  | **3** | **.656** | **.891** | **786.733** | **.657** | **.027** |
| | | | | ***Observation count*** | | | |
| *null* | | | - | .691 | 781.502 | .275 | .416 |
| ***final*** | **11.077*** | **3** | **.226** | **.741** | **763.103** | **.422** | **.243** |
| | | | | ***Motive/Effect count*** | | | |
| *null* | | | - | ≈ 0 | 524.464 | - | ≈ 0 |
| ***final*** | **34.417***** | **3** | **.293** | **.360** | **492.927** | **-** | **.095** |
| | | | | ***Total count of general annotations*** | | | |
| *null* | | | - | .707 | 638.834 | .286 | .421 |
| ***final*** | **14.220**** | **3** | **.338** | **.846** | **624.837** | **.675** | **.092** |

Note: In case of Motive/Effect reflective state, models were built with *student* as a random effect, while other models included *student nested within a course* as a random effect specification.

The likelihood ratio test between the *null* and *final* model for the ***Goal count*** as the outcome variable revealed a significantly better fit of the model that accounted for instructional conditions and prior experience in the tool use (Table 12). The model (Figure 12a) showed a significant effect of *instructional condition,* $F(1, 2.141) = 89.884$, $p = .009$, *experience*, $F(1, 11.377) = 30.347$, $p < .001$, and the interaction between *instructional condition* and *experience*, $F(1, 11.377) = 6.829$, $p = .023$. It seems that students in a graded condition and with the previous experience with the tool use tend to produce more goals in their reflections. Furthermore, pairwise comparison of least square means showed significant differences between the two instructional conditions: graded vs. non graded – $t(2.10) = 9.48$, $p = .009$, 95% CI [20.80, 51.60], as well as between the two levels of previous experience: no experience vs. with experience - $t(11.40) = -5.51$, $p < .001$, 95% CI [-26.40, -11.40].

Comparing the *null* and the *final* models for ***Observation count*** as a dependent variable, we again observed a better fit for the model that accounted for fixed and random effects (Table 12). However, the model (Figure 12b) revealed a significant effect of the previous experience only – $F(1, 71.996) = 7.678$, $p = .007$. This result could suggest that students with the previous experience in the tool use tend to produce more *observations* when reflect on their performance. Further analysis of pairwise comparison of least square means revealed a significant difference between the two levels of previous experience: no experience vs. with experience – $t(11.40) = -5.51$, $p < .001$, 95% CI [-26.40, -11.40].

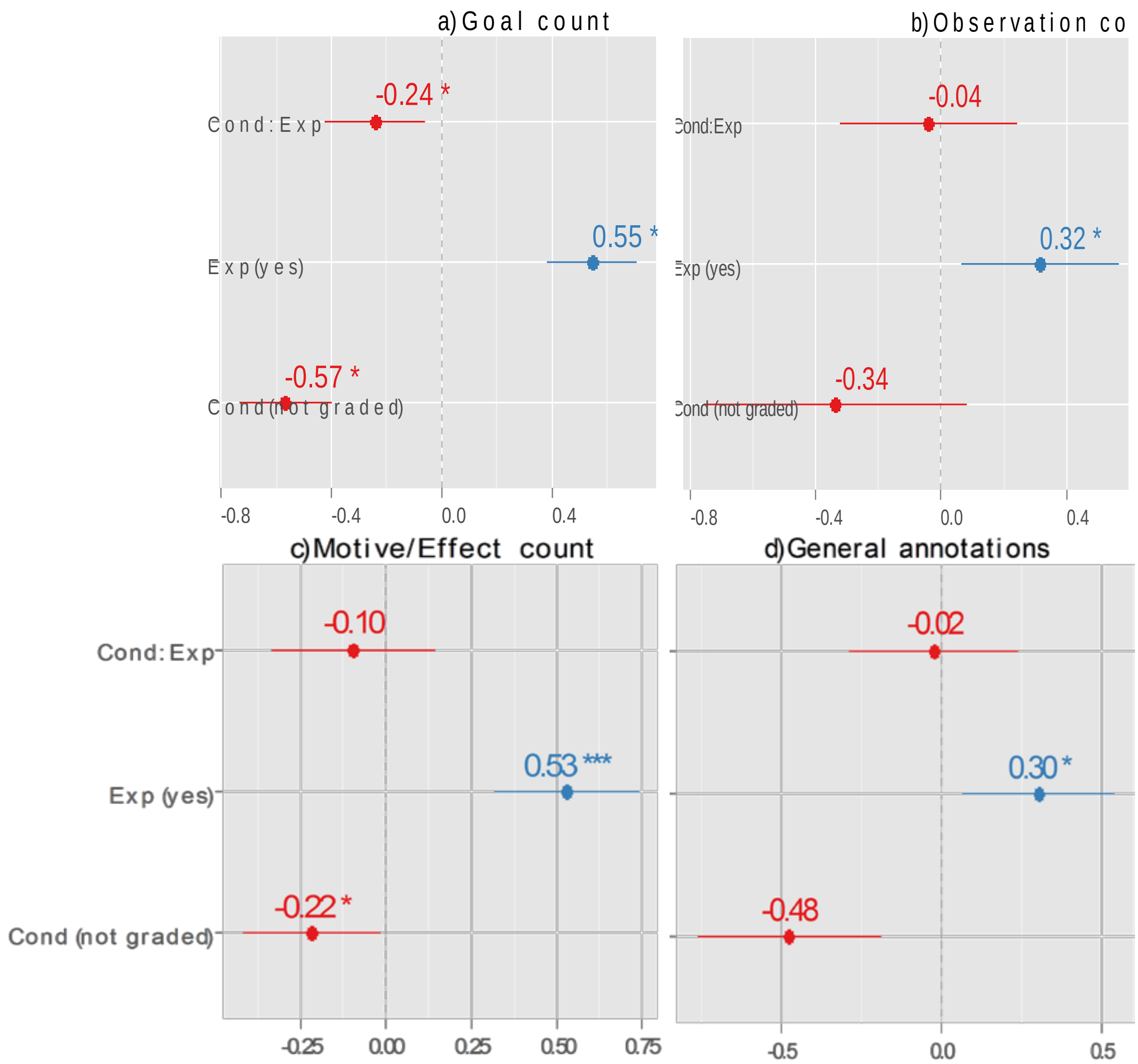



**Figure 12. Solutions for fixed effects (instructional condition and experience) on the specificity of reflective activities with video annotation**

Likewise, the previous models, the likelihood ratio test between the *null* and *final* model for the ***Motive/Effect count*** revealed a significantly better fit of the model that accounted for instructional conditions and prior experience in the tool use (Table 12). The model (Figure 12c) showed a significant effect of *condition*, $F(1, 87.161) = 9.697$, $p = .002$ and *experience*, $F(1, 8.338) = 27.408$, $p < .001$. The results might indicate that students with previous experience in the tool use and within the course with graded condition tend to produce more Motive/Effect reflections. The pairwise comparison of least square means further confirmed differences between the two levels of the previous experience: no experience vs. with experience - $t(88.30) = -5.24$, $p < .001$, 95% CI [-4.72, -2.12], as well as between the two instructional conditions: graded vs. non graded – $t(87.20) = 3.11$, $p = .002$, 95% CI [0.73, 3.33].

# Appendix E

***Example of activity design in case study 2***

Figure 13 shows an example of the activity scaffolding the use of CLAS. The example covers the topics of encoding numbers in base 2, 8 and hexadecimal, and provides as resources the video summarising the concepts and a PDF file capturing the material developed in the video. The work plan proposed for the activity requires:

1. Watch the video in CLAS
2. Write a summary of the video in the *general comments* area of the CLAS software. This comment will be visible to all students.
3. Place **three annotations** in the location in the video where a set of concepts are introduced. In **Error! Reference source not found.**, these concepts are the weight of bits in a sequence, the concept of octal representation, and the digits used in hexadecimal notation.

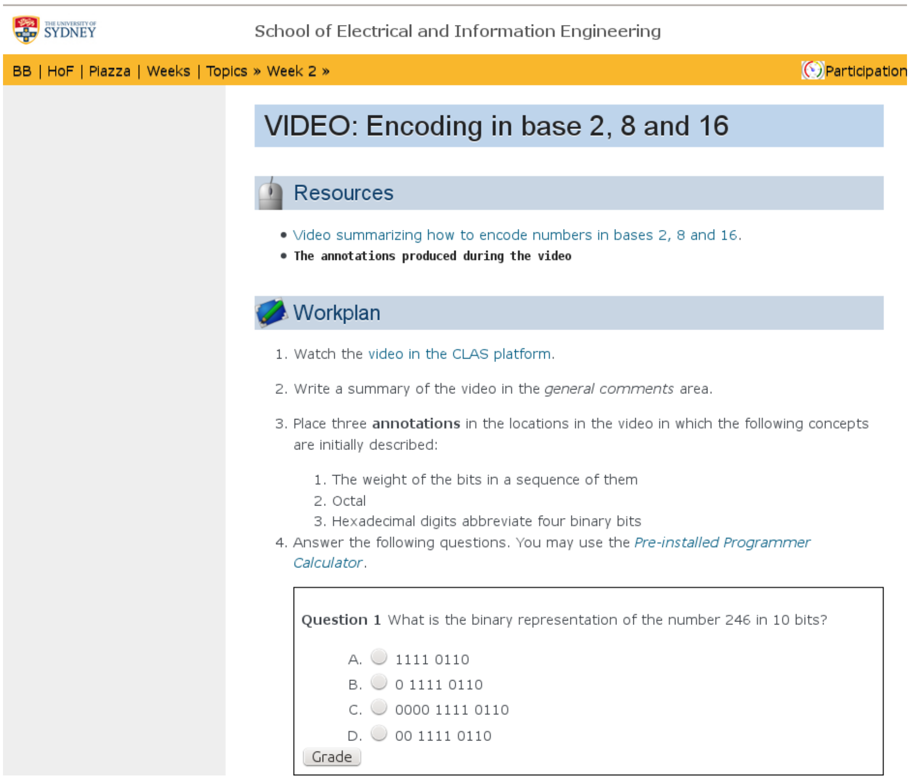



**Figure 13 Example of Activity using CLAS**

When students click on either the link provided as the first resource, or on the link in the first step of the work plan, the page shown in Figure 14 is loaded. The figure contains an actual capture of the material provided to the students in the case study. The top right corner shows the general comments, and the bottom left space shows the marks entered in the video to mark the concepts required in the instructions.

The general comments were made public by the instructors as a way of exposing students to other solutions and thus foster their judgement of learning. Analogously, all students could see the marks introduced in the timed annotations, so they could compare results with their peers.

The functionality in CLAS to introduce general comments was modified from that described in Section 2.1.1 to include a simple feedback mechanism. After the students write their general comments and click on the *Submit* button, an additional screen is presented on top of the general comments area as shown in Figure 15. The screen contains a set of points that the instructors previously defined as required and the students are asked if such points where included in the summary. Independently of the answer provided, the general comments are stored, and students are allowed to re-submit an unlimited number of times.

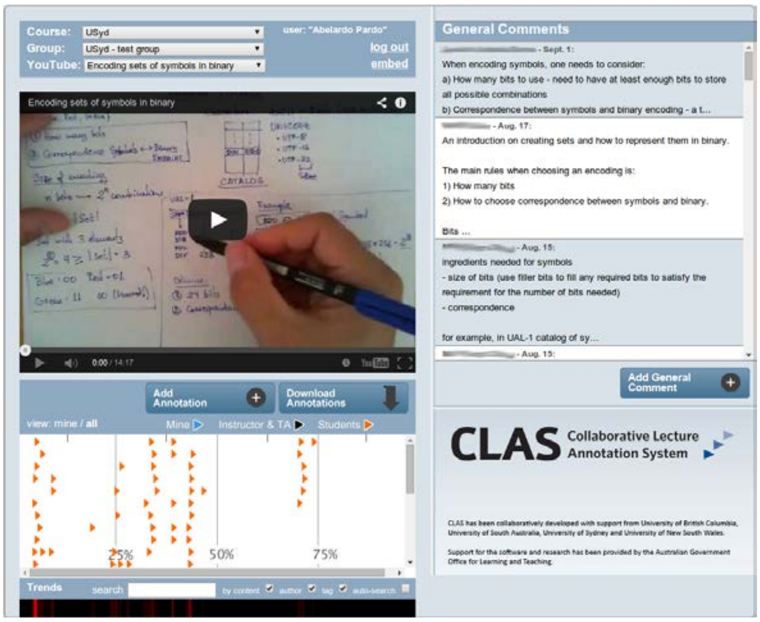



**Figure 14 Video activity in CLAS with General Comments and Annotations**

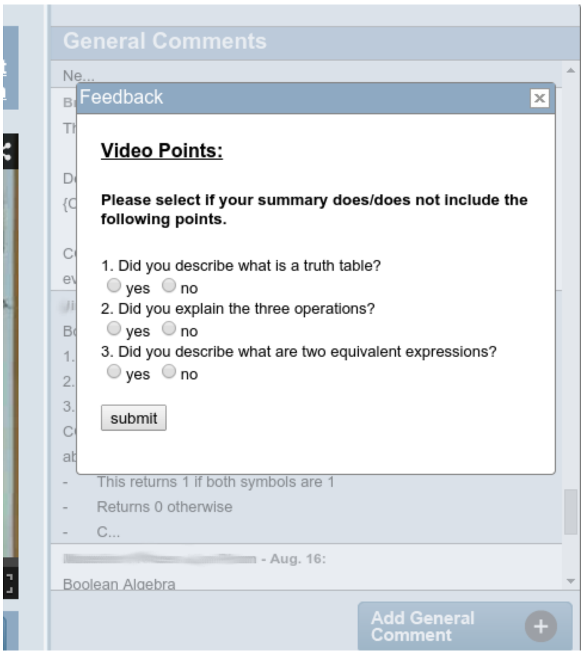



**Figure 15 Dialogue presented upon general comment submission**

The objective of this interaction is to provide students simple feedback about their comments and allow them to modify their entry. Moreover, based on the findings in the SRL research this feedback was introduced to assist students in increasing the accuracy of their judgement of learning (JOL).

# Appendix F

*Evaluation Reflections: Using Video Annotation Software to Develop Student Self-Regulated Learning (University of South Australia 2015)*

The aim of this project was to investigate the application of video annotations in higher education courses for the improvement of student self-regulated learning skills. A case study methodology was utilised to cover a range of disciplines and pedagogical contexts. The project's success was dependent on re-development of the video annotation platform CLAS – Collaborative Lecture Annotation System, which was originally co-developed by members of this project team and other international researchers.

**Objectives**

The project objectives were to:

- **a)** Investigate learner and teacher interactions with recorded video content
- **b)** Establish lead indictors for assessing student self-regulated learning
- **c)** Develop and establish data visualisations and student feedback processes
- **d)** Develop analytics and visualisations for staff to assist in curriculum decision making.

The guiding focus of the evaluation was to determine if the project's aims were achieved, and outcomes delivered, within budget and on time.

**Evidence**

This was a diverse project team lead by Professor Shane Dawson from the University of South Australia with partner institutions The University of Sydney, The University of New South Wales, University of Edinburgh and University of British Columbia. All team members have strong working relationships from former and other ongoing research projects. Furthermore there was a strong national and international reference group that provided input and advice throughout the life of the project.

In order to identify that the project's aims were achieved and outcomes delivered both formative and summative evaluation strategies were utilised. The external evaluator was provided with access to the key documentation from the project team and included in key project team communications.  In addition, the evaluator was a participant in virtual and face to face project team and reference group meetings. Throughout the lifecycle of the project the evaluator provided input and support via individual and group meetings with the two case study leaders in Australia – Drs Negin Mirriahi and Abelardo Pardo as well as a number of individual meetings with the Project Leader.

The Evaluator found several key factors that contributed to the successful achievement of the project aim and goals. These factors include:

- Regular meetings between the Project Leader and Evaluator were well supported by project plan updates and reports on activities. This ensured that the team were provided formative feedback to further enhance the proposed project outcomes.
- Strong project management, as demonstrated by extensive and appropriate documentation and insightful input to the project from the entire project team.
- Appropriate knowledge of institutional structures and priorities, ensuring that the activities undertaken related to institutional strategies and requirements in this emerging field.
- Diversity of skill sets in the project team ensured a range of perspectives and analytic techniques were represented. Diversity of ideas and novel approaches to the achieving the outcomes were promoted and supported.

**Project Management**

It has been documented that effective project management has the following elements:

- Identifying requirements,
- Establishing clear and achievable outcomes,
- Balancing the competing demands for quality, scope, time and cost,
- Managing the expectations of various stakeholders, and
- Adapting plans to overcome challenges.

From a Project Management perspective, the project was well managed and all stakeholder groups were involved. There was significant communication amongst all team members with involvement from the reference group assisting with project execution and promulgation of outcomes. From the outset it was evident that this was a well-led project with clear project goals and strategies. Inclusion from all stakeholders was facilitated by the project leader in a positive and generative manner. As this team was comprised of individuals who had previously worked together, team dynamics were not only strong but also key factors in the success of this project.

**Achievement of Outcomes**

The final report for the project provides a series of recommended actions for practice in how to enhance audio visual course content to assist students in developing greater self-reflective skills. While the case studies are diverse findings from all of them contribute broadly to the sector. A rigorous set of statistical analyses undertaken are detailed in the appendices for those interested in delving deeper. While, at a high level the findings are not overly surprising they do offer substantial insights into the use of video and video annotation for online and blended learning. For example, in order to improve student interaction and engagement with the technology there must be appropriate instructional scaffolds. However, as the case studies demonstrated when these scaffolds are well considered and embedded into curriculum practice student use of the tool not only increases but is applied in a manner that develops and improves student self-regulated learning proficiency in the longer term.

The final report provides an excellent overview of the three case studies and succinctly highlights recommendations for practice. This project has assisted in the re-development and re-branding of an open source tool for improving students self-regulated learning.